

An Error-Sampled Sweep-Position Control System

Abstract: This paper illustrates the application of sampled-data theory to the synthesis of a feedback control system using forward-path digital compensation to achieve high accuracy and fast response. It also presents a detailed description of the design and operation of a transistor system prototype. Developed for the specific purpose of controlling the position of an instantaneous portion of a cathode-ray-tube trace, the system has a number of unusual characteristics such as transport lag, digital integration, and numerous sampling rates. The practical manner in which these characteristics are implemented and the mathematical techniques used for system analysis should prove applicable to a wide variety of control-system problems.

Contents

1. Introduction	14
2. System Requirements	14
3. System Synthesis	16
Mathematical Formulation	
Frequency-Locus Analysis	
Root-Locus Analysis	
Predicted Performance	
4. System Operation	27
Discussion of Logical Block Diagram	
Error Detector and Sampler	
Analog-to-Digital Converter	
Storage and Transfer Circuits	
Digital-to-Analog Converter and Compensator	
System Performance	
5. Conclusions	34

1. Introduction

Considerable effort has been expended in recent years on the development of sampled-data systems. The major advantage of sampled-data or digital control is that, in principle, it permits us to make a continuous system behave exactly as desired at the sampling instants, that is, without overshoot or undershoot. The sampled-data design procedures are also important because they indicate methods for synthesizing digital process-control systems. Applications are also found in those systems where the data flow exists only at discrete instants of time, as in radar display systems. The application of sampled-data theory to the development of such a system is described.

As a consequence of the interest in this field, the familiar servomechanism theory of continuous-feedback systems has been extended to provide techniques for the

analysis and synthesis of sampled-data systems. The z -transform (a modified Laplace transform) as employed in the present synthesis is an outgrowth of work in this field by Barker,¹ Linvill,² Ragazzini and Zadeh,³ and Truxal.⁴ Through the application of the z -transform and other techniques, these investigators have found that much of the required system compensation can be accomplished while information is in digital form. In addition, many operations difficult to perform continuously can be carried out digitally with ease and accuracy.

The objective of this paper is first to demonstrate the application of z -transform techniques for sampled-data analysis to a particular system and then to illustrate the practical embodiment of the system in transistor circuits. In Section 2, which follows, system functions are described in detail. Section 3 describes the requisite mathematical analysis and development. Section 4 presents the physical system and describes its actual performance. This last section is so written that readers who are interested primarily in system implementation can go directly to it from Section 2 without loss of continuity.

2. System requirements

The system under consideration was designed to control the position of an instantaneous portion of a cathode-ray-tube sweep with sufficient precision to permit accurate measurements with its trace. Because the visible trace must not be distorted, it is necessary to delay correction of error until sweep recovery time. The scope of the problem is illustrated by Fig. 1a, which shows a sequence of sweep voltages applied to an oscilloscope deflection coil.

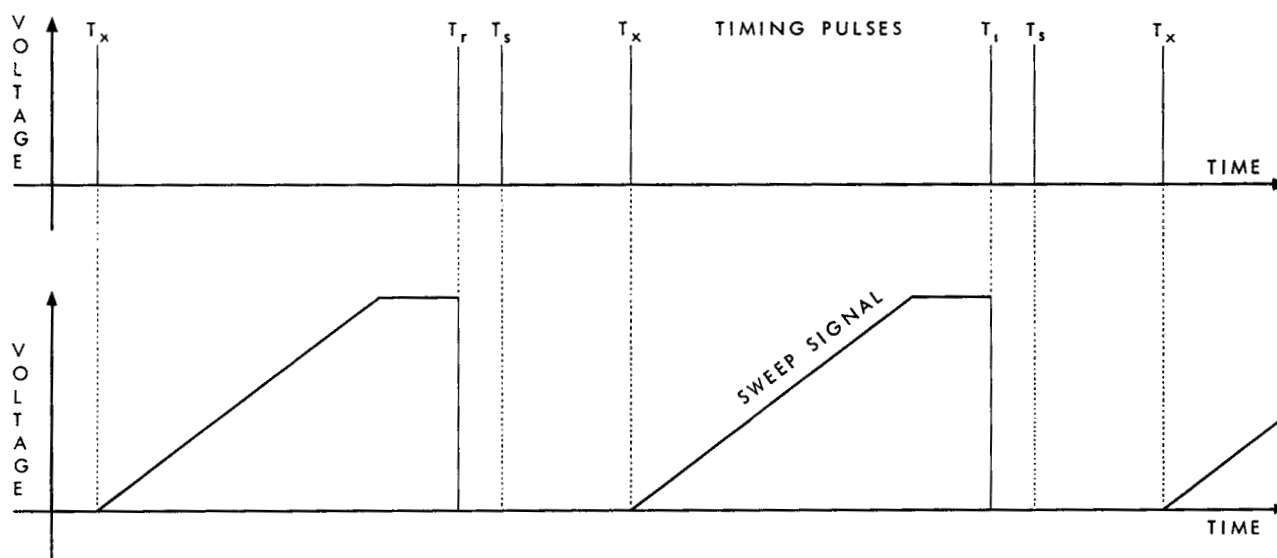


Figure 1a Sequence of sweep signals and timing pulses.

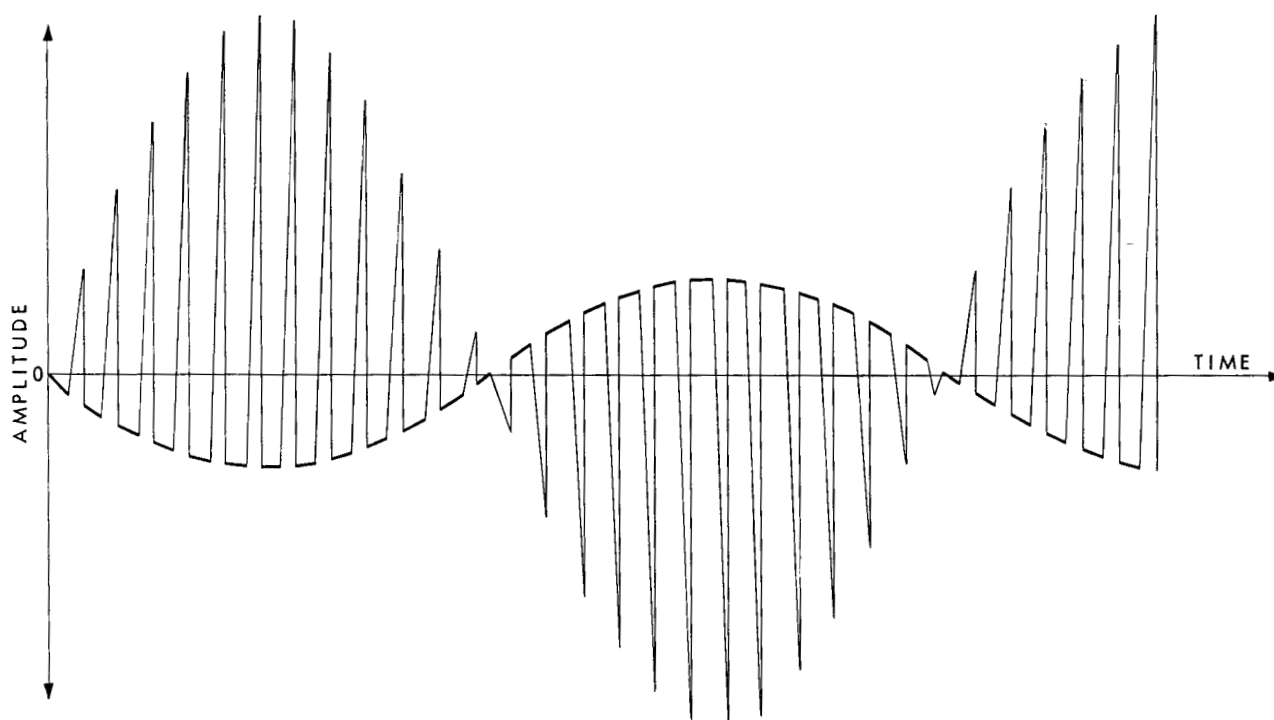


Figure 1b Resolved output over many sweeps.

T_s represents the system trigger pulse, that is, the beginning of each sweep period. T_r indicates the initiation of the sweep recovery period, and T_x represents the instant at which the sweep amplitude is to be controlled. Consider, for example, a plan-position-indicator (ppi) display in which this sweep voltage is ac-coupled to the scope deflection coils by means of a resolver. Because of the ac coupling, the deflection coil signal has no dc component, and hence its average amplitude over a cycle is zero.

Over many sweeps the resolver output appears as in Fig. 1b. It is obvious that the variations in the signal applied to the deflection coil between sweeps cause each sweep to start from a different point on the tube screen. To insure that each sweep originates from precisely the same point, the portion of the sweep just preceding the rise, or unblanking time, must be controlled. That is, T_x in Fig. 1a occurs at a point just preceding sweep initiation.

In the sector-scan form of presentation where only an amplified section of the ppi is displayed, the sweep

origin moves from the geometric center to one side of the tube and may effectively move several radii off the tube face. During this mode of operation, it is sometimes desirable to control that portion of the trace which describes an arc passing through the center of the tube. In this case the instant at which the sweep amplitude must be controlled (T_x) occurs at a point along the sawtooth. The exact location of T_x therefore depends on how far off center the sweep effectively originates.

For the system under discussion it will be assumed that pulses occurring at T_s , T_r , and T_x are available together with the inputs that serve as references during the sector-scan mode.

The fact that the error is available only at a precise instant on the sweep trace being controlled makes the need for an error-sampled system apparent. Because error correction must be delayed until sweep recovery time, T_r , there is also a need for some type of error storage. Furthermore, since the point T_x is a variable depending on the distance off center of the presentation, the delay time, $T_d = (T_r - T_x)$, must be continuously variable. Briefly, then, the error must be computed at time, T_x , stored until sweep recovery time, T_r , and then made available to effect correction.

The quantitative requirement placed on the system in this particular application is that any selected point on the trace should be correctly positioned to within 1/1000 of the tube diameter. The term "unit error" will therefore be defined as 1/1000 of the tube diameter, or the voltage or current at any specified point in the system necessary to move the sweep trace 1/1000 of the tube diameter. This precision must be maintained over a range of sweep frequencies from 200 to 1600 cps, corresponding to sweep periods of 600 to 5000 μ sec. For the sweeps to be used, it has been specified that the average sweep per error will be 4 units and the maximum will be 10 units. The system has therefore been designed to correct for a maximum of 10 error units per sampling period. In operation the point T_x may be shifted anywhere from within 10 μ sec after T_s to within 124 μ sec of T_r . The interval ($T_r - T_s$) is equal to 80% of total sweep time and thus may vary from about 500 μ sec for the fastest sweep to 4 msec for the slowest sweep. The required delay ($T_r - T_x$), may thus fall between the limits of 124 μ sec and 4 msec.

The most practical means by far of obtaining long and variable time delays is digital storage. Its use, however, necessitates converting the analog error pulse, or sample, to digital form and reconvert when it is desired to correct for the error, or to terminate delay. Although the complexity introduced by the analog-to-digital and digital-to-analog conversions may seem an unreasonable price to pay for digital storage, no other sufficiently flexible storage technique seems available at this time. There are, moreover, some compensating advantages. First, the digital-to-analog converter can itself provide several functions required by the system. Second, with error in digital form, system accuracy can be readily controlled.

As will be pointed out in the mathematical treatment, an integrating element is required in the forward control

path to provide required steady-state accuracy. The method of digital-to-analog conversion used readily provides this integration through its faculty of remembering all errors applied in the past. At the same time it serves as a low-pass filter necessary in error-sampled systems for removing high-frequency components of the error sample.

The maximum expected error of 10 units per sampling period is a mitigating factor in the design of the analog-to-digital converter. To achieve a quantizing accuracy within a limit of 10%, the converter need quantize to only 10 levels. Thus, conversion of the error to digital form not only provides variable time delay, integration and filtering, but also a straightforward means of controlling system accuracy.

Following is a summary of all the functions that must be included in the present system.

- a) Error detection.
- b) Error sampling.
- c) Analog-to-digital conversion of error sample.
- d) Error storage in digital form.
- e) Digital-to-analog conversion of error and integration with all past errors.
- f) Error correction.

3. System synthesis

• Mathematical formulation

Figure 2 is a block diagram of an error-sampled feedback control system that includes all the functions listed in the previous section. It should be stated at the outset that the system will be subjected mainly to ramp and step-function type inputs and disturbances. It is not difficult to show that under these conditions integral compensation within the loop is required. Consider a closed-loop feedback control system with a unit ramp input $r(t) = t$. According to z -transform analysis, $R(z) = Tz/(z-1)^2$ and the z -transform of the error signal, $E(z)$, is given by

$$E(z) = \frac{R(z)}{1+G(z)} = \left[\frac{Tz}{(z-1)^2} \right] \left[\frac{1}{1+G(z)} \right].$$

The final value theorem is used to find the steady-state error. Thus

$$e_{ss} = \frac{T}{\lim_{z \rightarrow 1} (z-1) \left[\frac{1}{1+G(z)} \right]} = \frac{T}{\lim_{z \rightarrow 1} [(z-1)G(z)]}.$$

The velocity error constant is defined as

$$K_v = \frac{1}{e_{ss}},$$

when $r(t) = t$. Therefore

$$K_v = \frac{1}{T} \left\{ \lim_{z \rightarrow 1} [(z-1)G(z)] \right\}.$$

In order that K_v be finite, $G(z)$ must contain a term of the form $[F(z)/(z-1)]$, where $F(z)$ does not contain any zeros at $(z=1)$. Inclusion of a term

$$[F(z)/(z-1)] = z/(z-1) \text{ in } G(z)$$

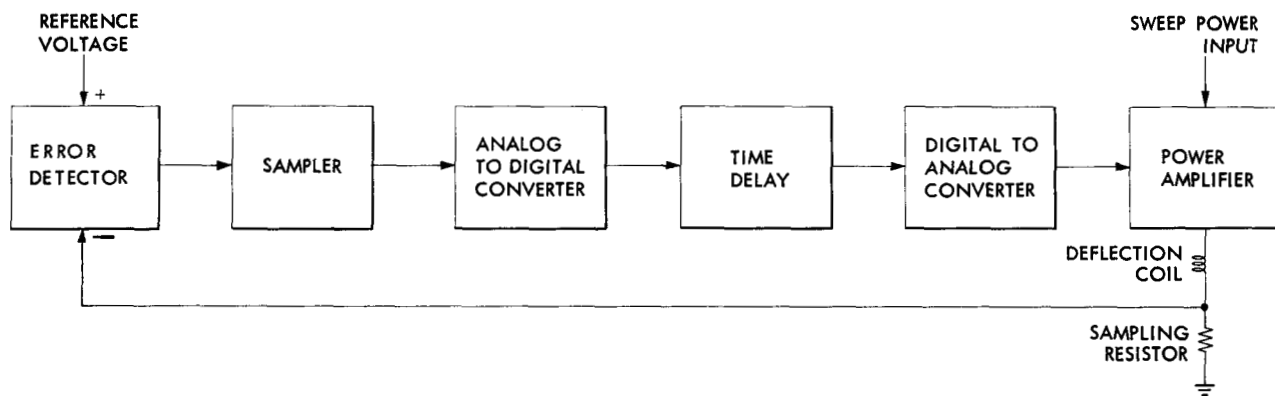


Figure 2 Block diagram of sweep-position control system.

corresponds to an integration in the s -plane and would permit proper closed-loop operation.

In view of the digital form of the internal loop signals made necessary by the time-delay requirements, a digital integration seems most expedient. Thus the functional block in Fig. 2 identified as the digital-to-analog converter includes circuits for performing the following operations on the input data:

1. Digital-to-analog pulse height conversion.
2. Algebraic summation of reproduced error sample with all past samples.
3. "Holding" or "clamping" action.

The clamping requirement stems from the fact that the output of the summation mechanism is a sequence of pulses, and it is mandatory that the amplitude of the output pulses be held or clamped in-between sampling instants. The clamping action in effect restores the power level in-between samples and forces the smoothing of the pulsed input data. Because of this smoothing action the transmission characteristics of the "clammer" can be regarded as similar to those of a low-pass filter. The combined effect of the algebraic summation and zero-order hold is to provide the required integration.

Having shown what type compensation will be required and how it is to be effected, we can proceed to formulate a complete mathematical representation. A sampled-data control system, such as the present one, that has deterministic input data, a specified sampling rate and compensation can be designed by either of the following techniques:

1. Frequency locus in the complex frequency plane.
2. Root locus in the z -plane.

In view of the fact that some care must be exercised in extending the techniques for continuous systems to sampled-data systems, both of these approaches will be utilized.

A third technique, which is based on a bilinear transformation from the complex z -plane to the complex w -plane is quite useful for systems that are compensated by stored-program digital computers.⁵ Since the present system is quite simply compensated, however, a decision was made for the direct z -transform methods.

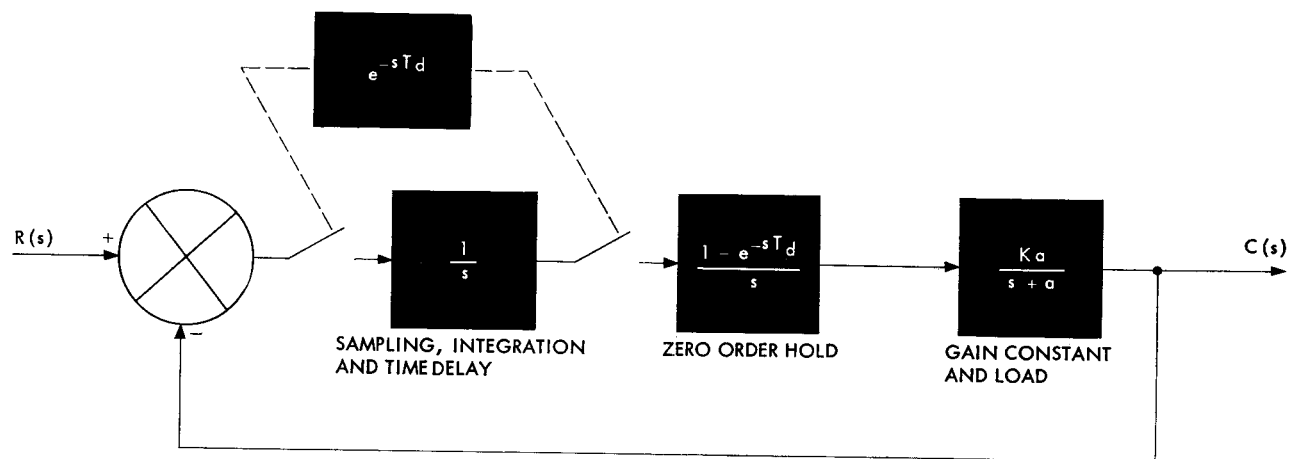
● Frequency-locus analysis

We begin the design with a consideration of the open-loop frequency locus (Nyquist diagram). We also assume that the M -Circle technique used for the analysis of continuous servo systems is applicable to error-sampled systems if the properties of the smoothing or hold circuits are such as to eliminate most of the high-frequency ripple which would otherwise appear at the load.⁶ There are two methods by which one can plot the frequency locus of a sampled network.⁷

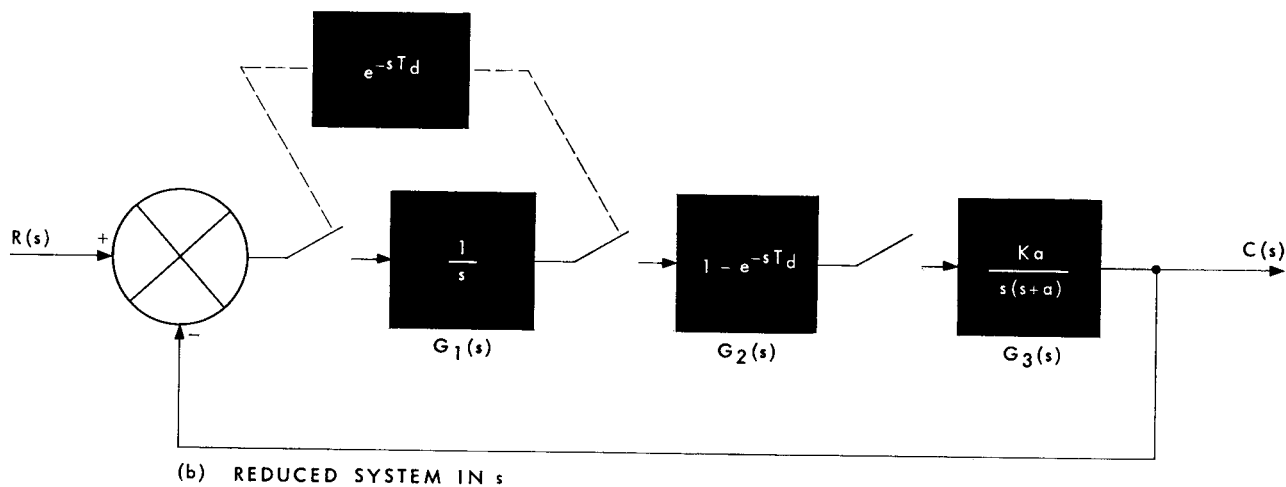
1. The loop transfer function is numerically evaluated in z for every frequency of interest. The computation is made by evaluating the absolute magnitude and phase of the polynomial in z as the phase angle of z is varied from 0 to 180° on the unit circle.
2. The transfer function in s of the sampled network is expanded as an infinite series and a sufficient number of terms are used to approximate the expression in s at each frequency of interest to the desired degree of accuracy.

If we apply the M -Circle assumptions to the system under consideration in this paper, the form of continuous network compensation, overshoot, frequency at which the system resonates, and required gain for the specified dynamic and steady-state performance can be evaluated from the frequency locus.⁸

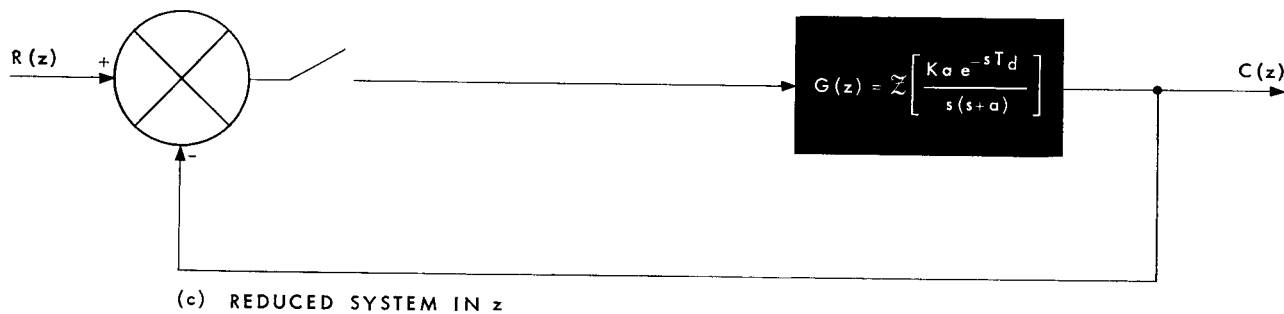
The first step in the synthesis procedure is the representation of the system as shown in Fig. 2 in terms of transfer functions in s , which in turn can be transformed into transfer functions in z .³ In the $\mathcal{Z}$ -transform diagram shown in Fig. 3a, the necessary integration and time delay is signified by inclusion of the e^{-sT_d}/s block and the $1/s$ block in the forward path of the loop. At this point we make the assumption that the analog-to-digital (and digital-to-analog) conversion process does not result in a significant error when the $1/s$ representation is used for the entire process. In effect we are stating that quantization errors introduced by the conversion and reconversion process do not contribute significantly to the behavior of the present system. Although this cannot be said of error-sampled systems in general, it is justified in the present case because the minimum quantization level is an ex-



(a) REQUIRED CONTROL SYSTEM IN $\mathcal{L}$ -TRANSFORM REPRESENTATION



(b) REDUCED SYSTEM IN s



(c) REDUCED SYSTEM IN z

18 Figure 3 System reduction procedure for obtaining open loop $G(z)$.

remely small portion of the total holding-circuit capacity. The remainder of the Laplace-transform-system block diagram is conventional except for the representation of the sampler. The latter has been adequately described in the literature⁸ and is also treated in detail in an IBM Research Report on the z -transform method of analysis for sampled-data systems.⁹

For purposes of reduction to the z -transform diagram (Fig. 3c), the sequence of operation of the digital-to-analog converter is initially represented in Fig. 3a as an integration followed by the necessary zero-order clamp, and reduced in Fig. 3b to the intermediate s -diagram. It should be noted that the zero-order clamp or hold can be described analytically by two step functions separated in time by the sampling period. The first step function is positive-going from zero to unity amplitude; the second is negative-going from unity amplitude to zero.

The open-loop transfer function, $G(z)$, is given by

$$G(z) = G_1(z)G_2(z)G_3(z);$$

However,

$$G_1(z)G_2(z) = \mathcal{Z}(e^{-sT_d}/s) \times \mathcal{Z}(1 - e^{-sT}) \\ = e^{-sT_d} \left(\frac{z}{z-1} \right) \left(\frac{z-1}{z} \right) = e^{-sT_d}.$$

Thus

$$G(z) = \mathcal{Z}[e^{-sT_d}G_3(s)] = \mathcal{Z}\left[\frac{Kae^{-sT_d}}{s(s+a)}\right],$$

and Fig. 3b may be reduced to Fig. 3c.

$G_3(s)e^{-sT_d}$ may be expanded into partial fractions and the z -transform obtained directly, giving

$$G(z) = K \left[\frac{z(1 - e^{-a\Delta T} - e^{-aT})}{(z-1)(z - e^{-aT})} \right],$$

where

$$z \triangleq e^{sT},$$

$$K \triangleq \text{Effective loop voltage gain,}$$

$$T = \text{Sampling period,}$$

$$T_d = \text{Transportation lag (time delay), and}$$

$$\Delta T = T - T_d.$$

Note that although the z -transform could have been taken immediately from the s -plane representation of Fig. 3a, the effect on the system $G(z)$ of the zero-order clamp would not be particularly apparent. Breaking down the loop into Figs. 3b and 3c reveals the internal structure of the system and so provides the designer with an increased understanding of the mathematical properties of the system components.

We are now ready to obtain the frequency locus and shall do so by the first of the two methods previously listed, that is, by evaluating $z \triangleq e^{sT}$ at all frequencies. The second method, that of expanding the s -transfer function as an infinite series, gives less accurate results and re-

quires almost as much effort. Using the method selected, we first define z as equal to $e^{j\omega T}$, and then move around a unit circle in the z -plane to evaluate the change in vector location of the poles and zeros of $G(z)$ with frequency. By virtue of this process the left half of the s -plane is mapped into the interior of the unit circle in the z -plane. Thus a test for stability for a sampled system is that no roots of the characteristic equation $1 + G(z) = 0$ be outside the unit circle in the complex z -plane. In terms of the frequency locus, the normal Nyquist criterion still holds, i.e., the locus is not permitted to encircle the $(-1, j0)$ point in the $j\omega$ -plane.

Figure 4 illustrates several outstanding differences between the frequency loci of the present system calculated for continuous and for error-sampled modes of operation, with identical time delay in each case. Continuous loop gain, K_c , is equal to $1/T$ times the loop gain, K_s , of the sampled system, i.e., $K_c = K_s/T$. Note that the locus of the continuous system spirals into the origin as expected, whereas the locus of the error-sampled system does not. This is due to the fact that the locus of the former is aperiodic, while the locus of the latter is periodic, with a period of $T_0 = T_s$. The frequency distributions for the two systems also differ.

A most interesting inference that can be drawn from Fig. 4 is that where time delay is required and velocity-error constants are similar, better performance may be achieved with an error-sampled system than with a continuous system. The possibility of improvement is suggested by the fact that the locus of the error-sampled system is further away from the $(-1, j0)$ point at the crossover frequency than the locus of the continuous system. The magnitude of this effect will, however, depend on the ratio of the sampling period to the time delay, and it cannot be stated as a general rule that an improvement is always possible.

Also studied was the effect on system performance of the effective-load time constant, which in this case is the load inductance divided by the sum of the amplifier output resistance and load resistance. The Nyquist diagram for the case with 90 ohms power-amplifier output impedance indicates a transient response whose settling time was the minimum required by the system specifications; however, the Nyquist diagram for the case with zero power-amplifier output impedance indicates a very poor dynamic behavior. The conclusion to be drawn is that to insure satisfactory performance, the effective-load time constant must be decreased by at least a factor of ten.

• Root-locus analysis

The root-locus technique, which permits ready design for a given damping ratio and damping factor, is in common use for continuous feedback systems and may be extended to sampled systems with little difficulty. By using the definition $z = e^{sT}$ and drawing the locus of the roots of the system characteristic equation, $1 + G(z) = 0$, contours of constant damping factor and damping ratio can be transformed from the s -plane to the z -plane.

In the continuous system all transient terms arise from

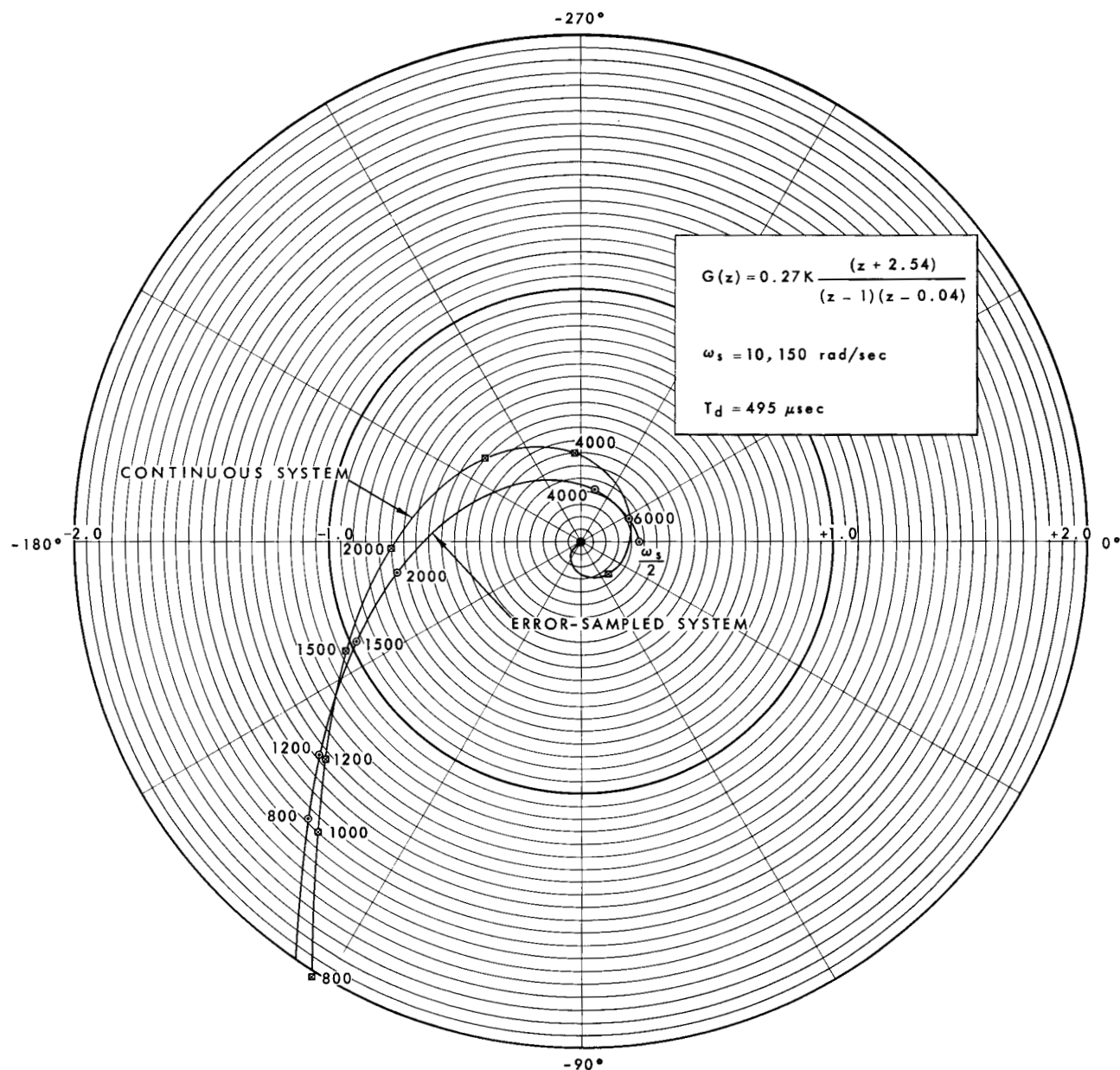


Figure 4 Frequency loci for identical error-sampled and continuous control systems, both with the same time delay. Gain $K=1$ in each case.

the general quadratic factor of the characteristic equation,
 $(s^2 + 2\omega_0\zeta s + \omega_0^2)$

and assume the form

$$Ke^{st} = Ke^{-\zeta\omega_0 t} e^{j\omega_0 t \sqrt{1-\zeta^2}} = K(e^{-\sigma T})(e^{j\omega_0 T \sqrt{1-\zeta^2}}),$$

where ζ is the damping ratio, σ the damping factor, and ω_0 the undamped natural frequency of the quadratic term. In the s -plane, contours of constant σ are lines parallel to the imaginary axis, and contours of constant ζ

are radial lines whose angle with the negative real axis is $\cos^{-1}\zeta$. If we let

$$z = e^{sT} = (e^{-\sigma T})(e^{j\omega_0 T \sqrt{1-\zeta^2}}),$$

these contours of constant damping ratio and damping factor may be translated to the z -plane. A number of the loci obtained in this manner are shown in Fig. 5 along with contours of constant $\omega_0 T$. Since these loci are symmetrical about the real axis, only the positive imaginary half of the z -plane is shown.

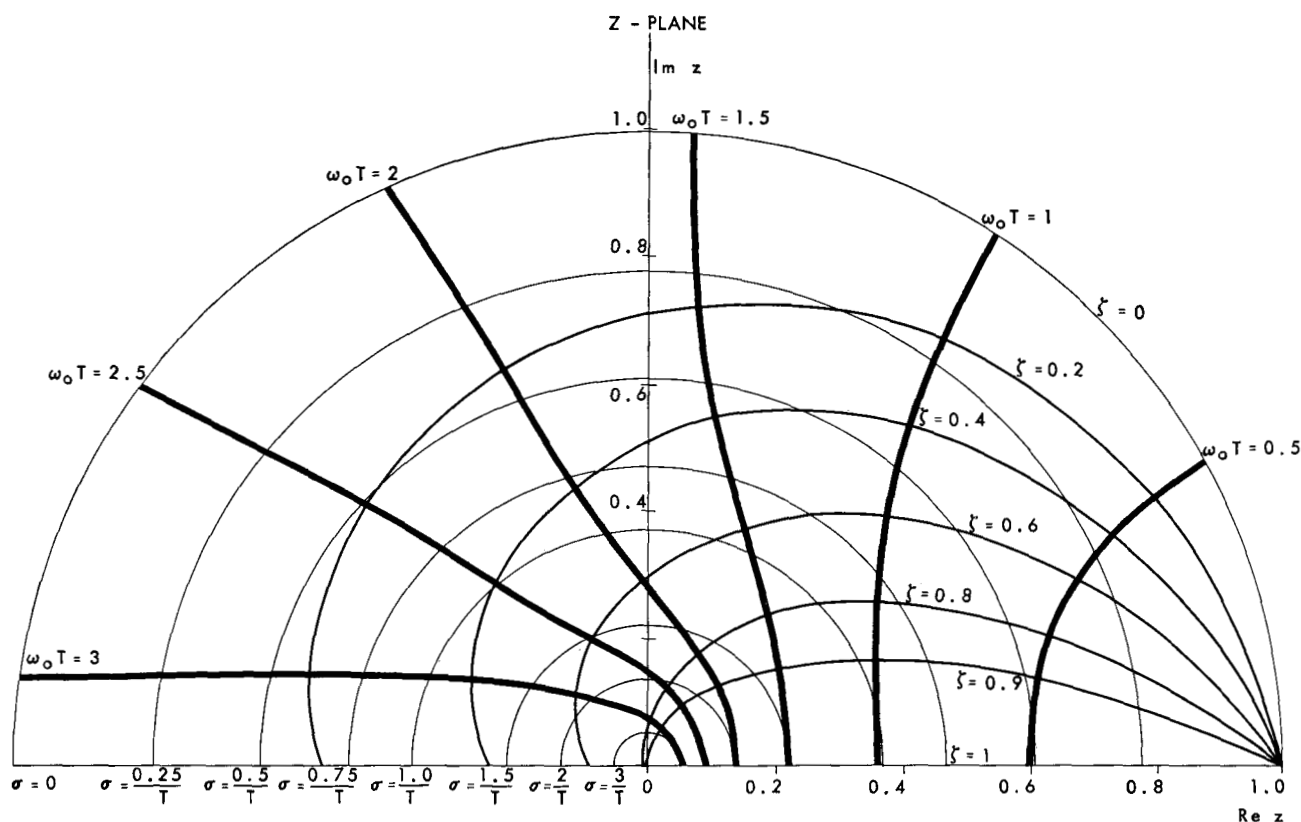


Figure 5 Contours of constant-damping factor σ , damping ratio ζ and $\omega_0 T$.

Once the location of the roots of the characteristic equation in the z -plane is known, the nature of the transient response of the system at sampling instants can be determined. Consider the case in the present system for the shortest sampling period (600 μ sec) and maximum delay (540 μ sec). Assuming that the output impedance of the dc power amplifier is zero (which gives the limiting value of the load time constant),

$$G(z) = K \left[\frac{z(1 - e^{-a\Delta T}) + e^{-a\Delta T} - e^{-aT}}{(z-1)(z - e^{-aT})} \right] = \frac{0.27K(z+2.5)}{(z-1)(z-0.05)}$$

where

$$a_{\min} = \frac{R_{\min}}{L_{\max}} = \frac{10}{2 \times 10^{-3}} = 5,000,$$

and

$$\Delta T = T - T_d = 600 - 540 = 60.$$

All these conditions are in the direction giving the poorest transient response. Thus the largest time constant, longest time delay, and shortest sampling period are used.

Obtaining the characteristic equation, we get

$$1 + G(z) = 0 = z^2 + z(0.27K - 1.05) + 0.05 + 0.68K.$$

Thus for $K=1$,

$$(z - 0.39 + j0.76)(z - 0.39 - j0.76) = 0.$$

The location of these roots on Fig. 5 gives $\zeta=0.15$, $\sigma T=0.15$, and $\omega_0 T=1.1$. The actual system response to a unit step is given by $1 - (e^{-\sigma nT} \cos n\omega T)$, where

$$n = 1, 2, 3, \dots,$$

$$\sigma T = 0.15, \text{ and}$$

$$\omega T = \omega_0 T \sqrt{1 - \zeta^2} = 1.1 \sqrt{1 - (0.15)^2} = 1.08.$$

Then

$$f = \frac{\omega}{2\pi} = \frac{1.08}{2\pi T} = \frac{1}{5.8T} \text{ cycles/sec}$$

and the response at sampling instants is thus given by

$$C(nT) = 1 - e^{0.15nT} \cos(2\pi n/5.8),$$

which is shown as the smooth curve in Fig. 6a. It is also possible to obtain this response through long division of $R(z)G(z)$ by $HG(z) + 1$. Doing so results in a series in descending powers of z , $(0.25z^{-1} + 1.15z^{-2} + 1.65z^{-3} + \dots + C_n z^{-n})$, the coefficients of which represent $C(t)$ at $t = nT$. It can be seen from Fig. 6a that the results for the two methods correspond quite well. It should be pointed out that the response obtained by either method is valid only at sampling instants, 0, T , $2T$, etc., since the unmodified z -transform is used throughout. In general, for

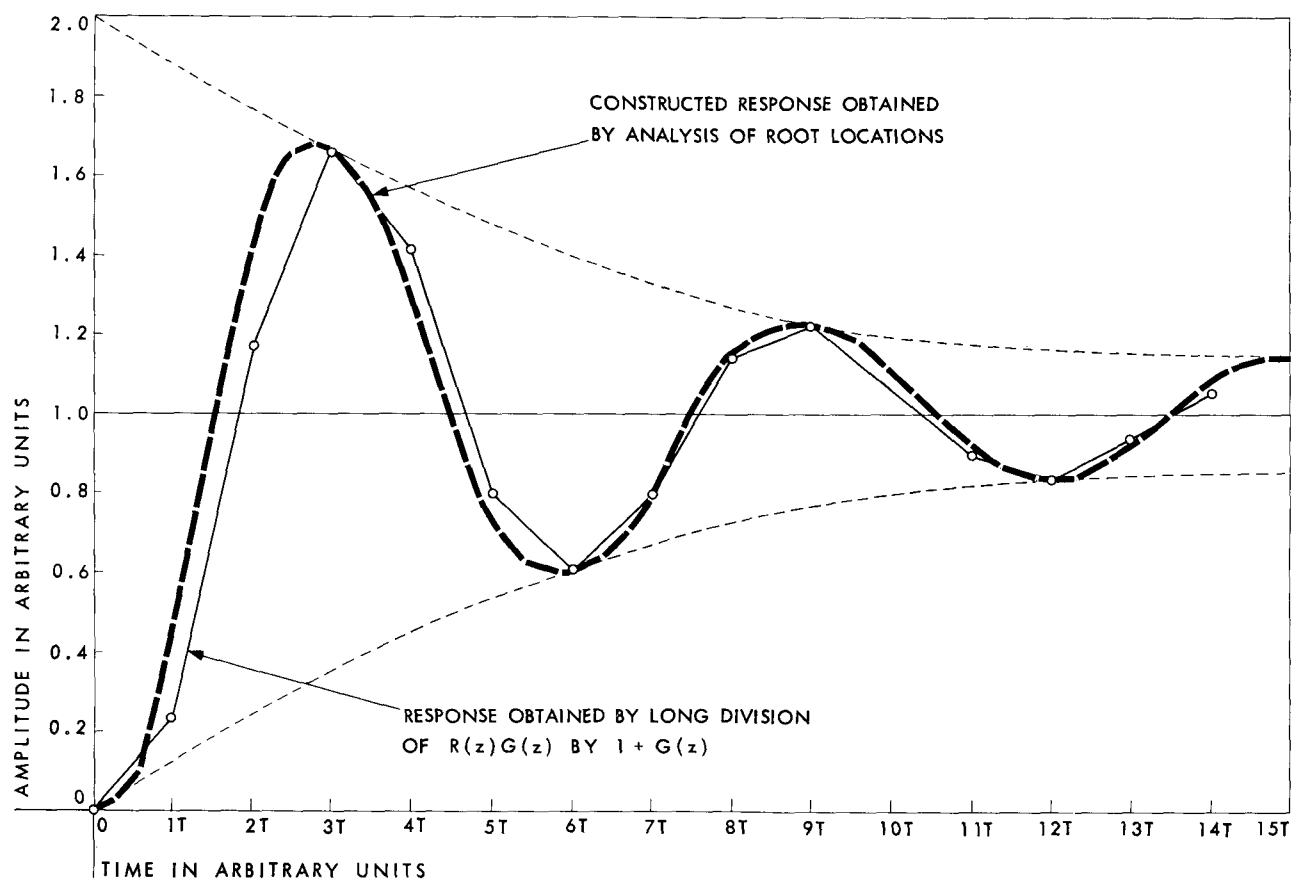


Figure 6a System response to unit step obtained by two independent methods for the case with zero amplifier output impedance.

systems having a higher than second-order characteristic equation, the exact transient response cannot be so readily obtained from the location of the roots in the z -plane. In practice, however, certain roots usually dominate and a good approximate response may be obtained without undue labor.

It is plain from Fig. 6a that, if transient response is to be sufficiently damped, amplifier-output impedance must be made greater than the zero value for which the curves were obtained. If a 70-ohm value is substituted, $a_{\min} = R/L_{\max} = 40,000$. Therefore

$$G(z) = \frac{0.92K(z+0.09)}{z(z-1)}.$$

If $K=1$,

$$1+G(z)=0=z^2-0.08z+0.09,$$

or

$$z=0.04 \pm j0.3.$$

The location of these roots on Fig. 5 gives $\zeta \approx 0.62$, $\sigma T \approx 1.25$, and $\omega_0 T \approx 2$, which correspond to a well damped response. The output curve constructed from these values for a step-function input is shown in Fig. 6b. The approximate actual response obtained by the impulse

response method, which is also shown, is in good agreement at sample times. The fact that the response obtained from the z -plane location of the roots of the characteristic equation is valid only at sampling instants is thus clearly brought out.

Considering the fact that the maximum amplitude of the step function to which the system will be subjected is 10 units, the transient response should be such that the system settles to within 1 error unit (or 10%) in one sample period. For a unit step then, the system should settle within limits of 0.9 to 1.1 in one sample period. In Fig. 6a the system requires over 15 sample periods to settle within satisfactory limits, while in Fig. 6b the system does so within one sample period. For satisfactory performance, then, the load time constant should be less than $1/40,000 = 25 \mu\text{sec}$. The bandwidth of the dc amplifier should therefore be greater than

$$1/25 \times 10^6 / 2\pi = 6,000 \text{ cps.}$$

The bandwidth must be greater than this in any case in order to pass the sweep waveform.

An insight into what happens as this load time constant decreases (or the output impedance increases) is demonstrated by the root locus of the system in the z -plane. For

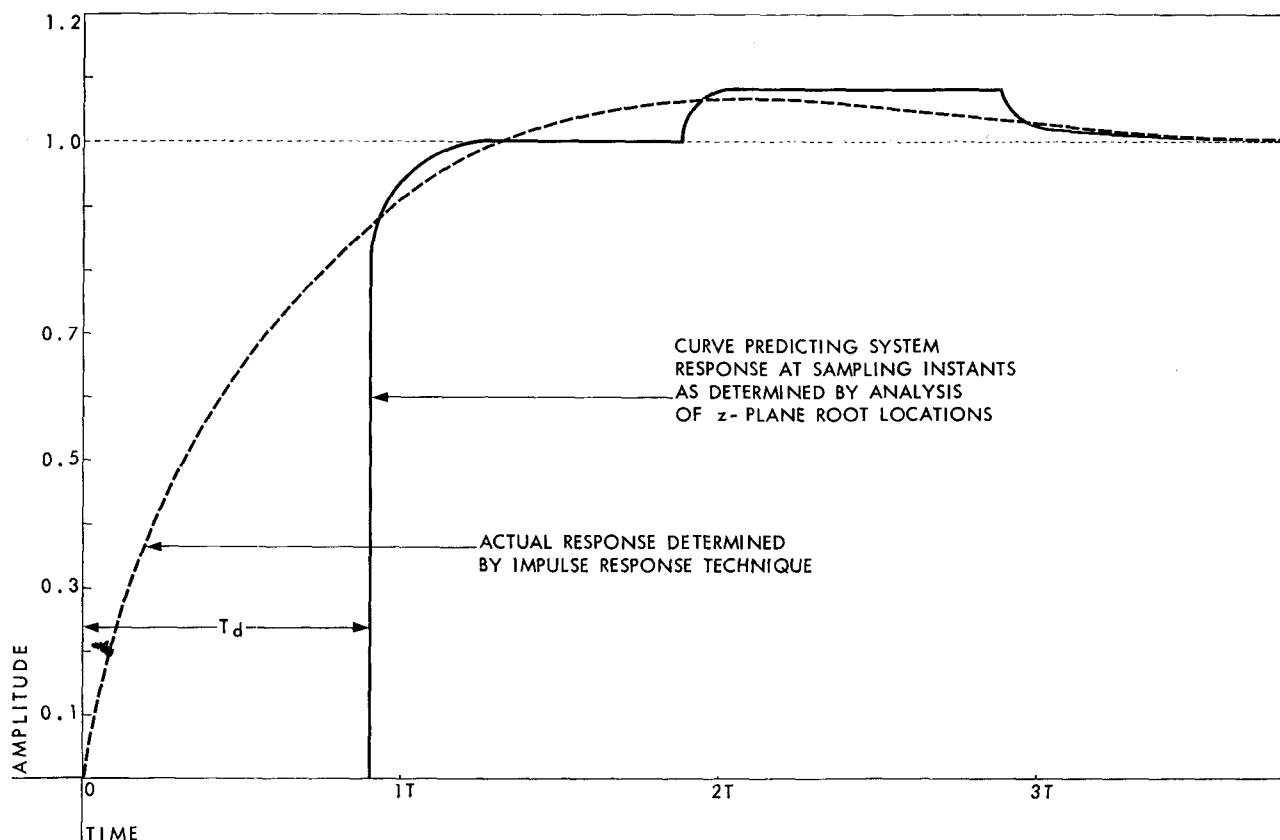


Figure 6b System response to a unit step comparing actual response with predicted response at sampling instants for the case with 70 ohms amplifier output impedance.

$$G(z) = 0.27K(z + 2.5)/(z - 1)(z - 0.05),$$

which corresponds to the poor response of Fig. 6a, the locus is as shown in Fig. 7a. The location of roots at $K=1$ corresponds to that previously determined. The best transient response occurs at a gain of about 0.5 or 0.6. The locus for $G(z) = 0.92K(z + 0.09)/z(z - 1)$ corresponding to the response of Fig. 6b is shown in Fig. 7b. The effect of decreasing the load time constant from 200 μsec to 25 μsec is to move the zero of $G(z)$ from -2.5 to about -0.1 , thereby improving the locus shape as shown. In this case the transient performance is excellent at gains near 1.0 and is quite good for gains of 0.6 to 2.0.

An interesting observation regarding application of root-locus techniques to sampled systems is that the sampled system with time delay presents no more difficulties than the same system without time delay. Such is not the case with continuous systems. Application of the root-locus method to a continuous system with transport lag is much more laborious than is its application to the same continuous system without transport lag.¹⁰

• Predicted performance

Figures 8 and 9 demonstrate and summarize the effects on system transient response of variations in sampling rate, time delay, and effective load time constant. The

responses shown were determined with the techniques described and as previously stated, are valid at sampling instants only. Fig. 8 illustrates the response to a unit step at the input and Fig. 9 the response to a unit-step disturbance at the load. As could be expected, the poorest transient performance is obtained in parts (a) where the conditions of longest relative time delay, fastest sampling rate, and largest load time constant are combined. Parts (b) indicate the large improvement effected by increasing the sample period to the maximum of 5000 μsec , and parts (c) demonstrate the improvement over parts (a) brought about by reducing the time delay from 540 to 180 μsec .

From the standpoint of sample period and time delay, then, the conditions of Figs. 8a and 9a represent the worst case which will be encountered in operation of the system (as far as transient performance is concerned). Recalling that the maximum error per sample can be only 10 units without saturation and that the system must correct to within one error unit in one sampling period, then for a unit step at the input the output should follow within 10% in less than one sample period after the error is sampled. Similarly for a unit-step disturbance at the load, there should be a reduction to less than 10% in less than one sample period after the disturbance is

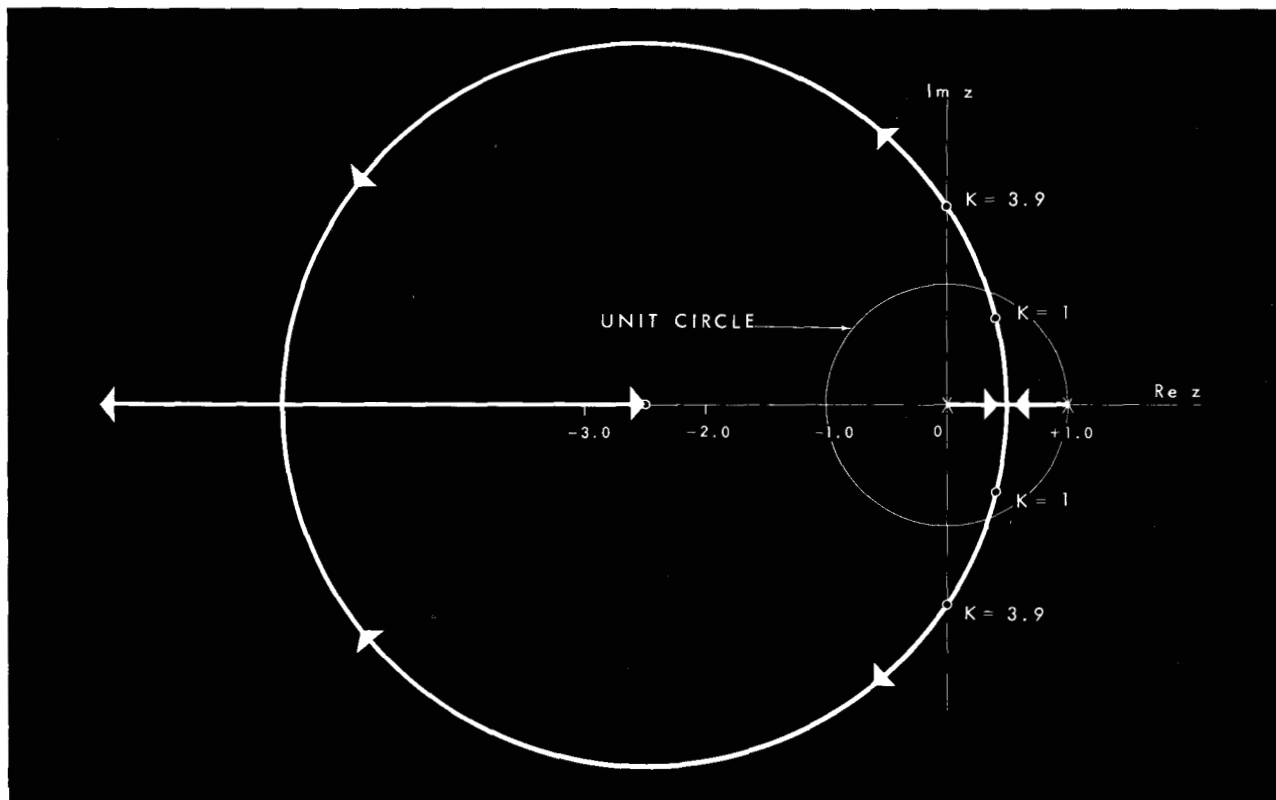


Figure 7a Root locus for $G(z) = \frac{0.27K(z+2.5)}{(z-1)(z-0.05)}$.

Figure 7b Root locus for $G(z) = \frac{0.92K(z+0.09)}{z(z-1)}$.

sampled. By effectively lowering the load time constant to 25 μ sec (either by raising the amplifier output impedance to 70 ohms or by using a lead network before the amplifier) the responses are improved to the extent that the above requirements are met as shown in Figs. 8d and 9d.

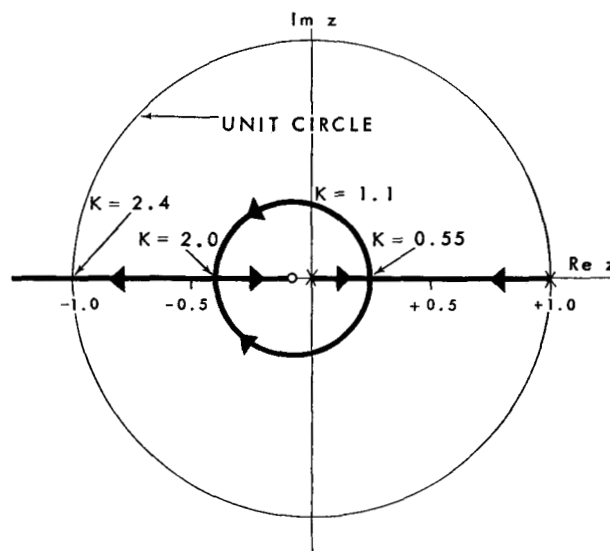
Figure 10 clearly demonstrates the system response to a unit ramp input and load disturbance, respectively. The disturbance in ppi operation and the input and disturbance in sector-scan operation is a low-frequency sine wave, $A \sin \omega t$, the z-transform of which is

$$Az \sin \omega T / z^2 - 2(\cos \omega T) + 1.$$

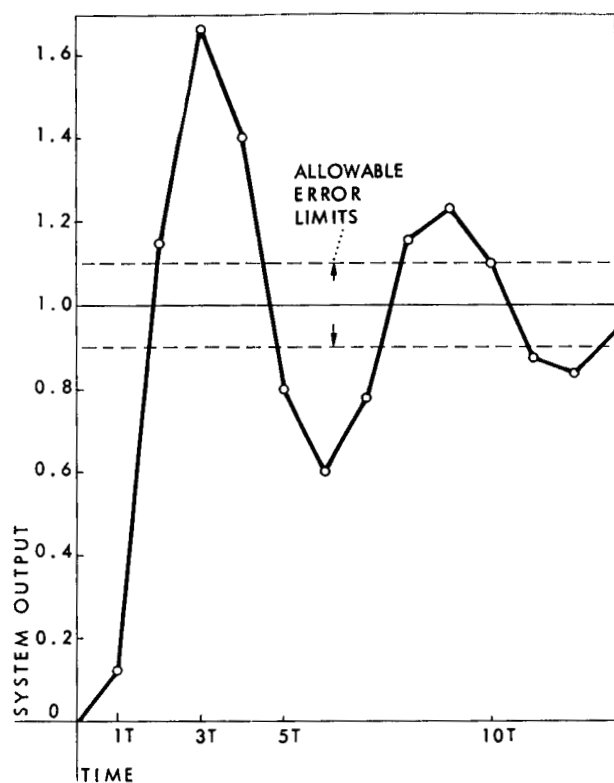
Since $T_{\max} = 5 \times 10^{-3}$ sec and $\omega_{\max} = 2\pi \times 1/2$ radians per sec (the maximum resolver rotation frequency), then $\omega T \leq \pi \times 5 \times 10^{-3}$, $\sin \omega T \approx \omega T$, and $\cos \omega T \approx 1$. The z-transform of $A \sin \omega t$ thus becomes approximately

$$A \omega \left[\frac{Tz}{(z-1)^2} \right],$$

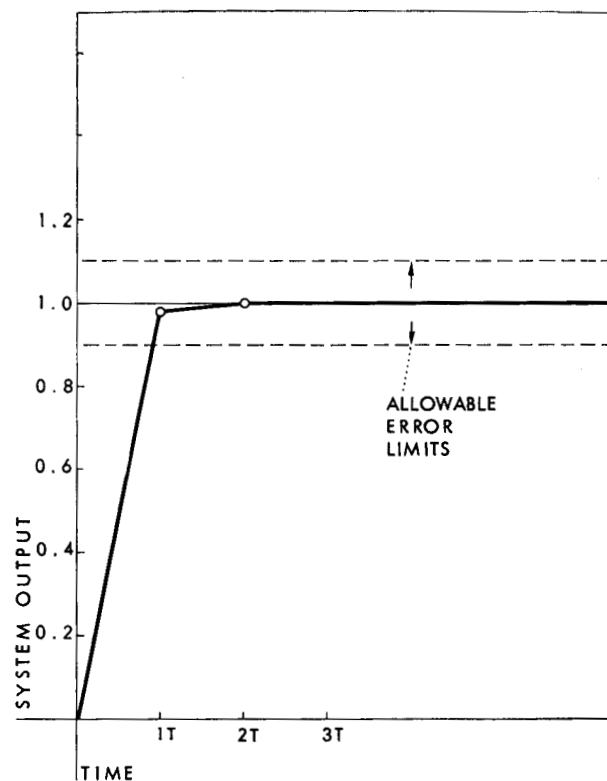
which is the z-transform of a ramp function, $A \omega t$. Figure 10 graphically indicates that the steady-state error in the presence of a unit ramp input or disturbance is T . In general it is T/K , K being 1 in this case. Since the maximum steady-state error then is $A \omega T$, the worst steady-state error is obtained with the slowest sampling rate. Here $\omega_{\max} = \pi \text{ rad/sec}$, $T_{\max} = 5 \times 10^{-3}$ sec, and $A_{\max}$



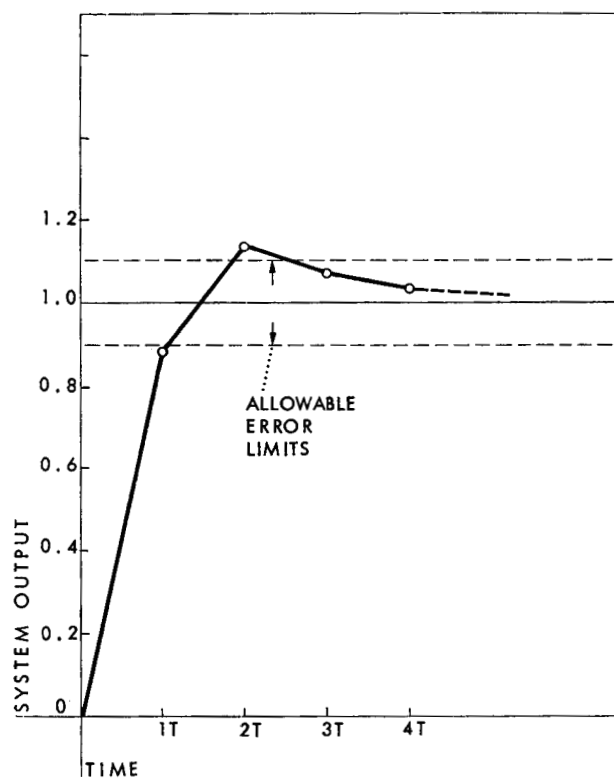
referred to sampled output is approximately 1.6 v. Thus $(A \omega T)_{\max} \approx 1.6 \times \pi \times 5 \times 10^{-3} = 24$ mv, which represents about a 1.5 unit error at the controlled output. Since this maximum would be present only under the most adverse conditions, it may be concluded that the system could readily be compensated to insure satisfactory performance.



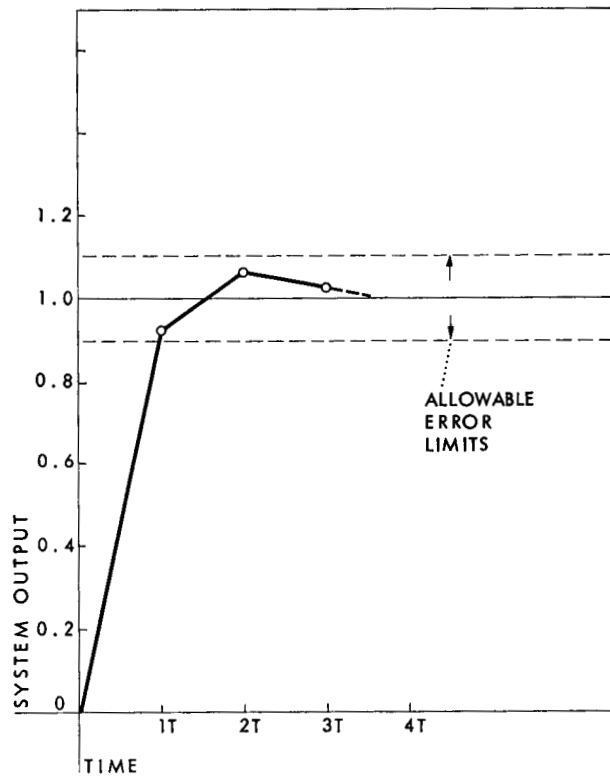
a) $T=600 \mu\text{sec}$, $T_d=540 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.



b) $T=5000 \mu\text{sec}$, $T_d=4000 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.

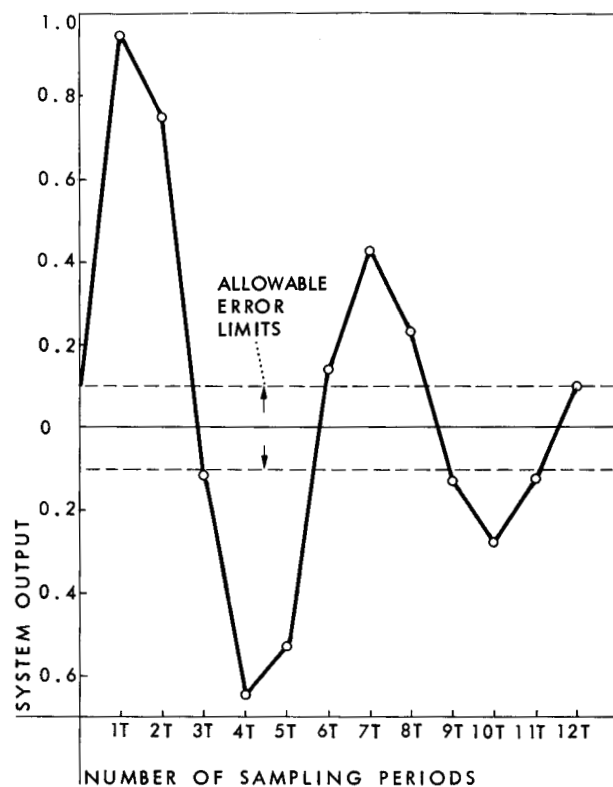


c) $T=600 \mu\text{sec}$, $T_d=180 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.

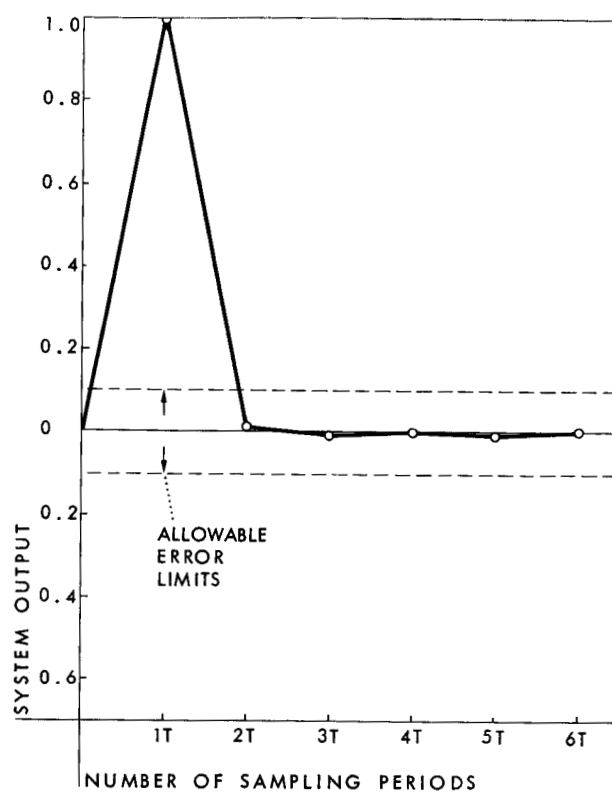


d) $T=600 \mu\text{sec}$, $T_d=540 \mu\text{sec}$, $K=1$, $a=R/L=40 \times 10^3$.

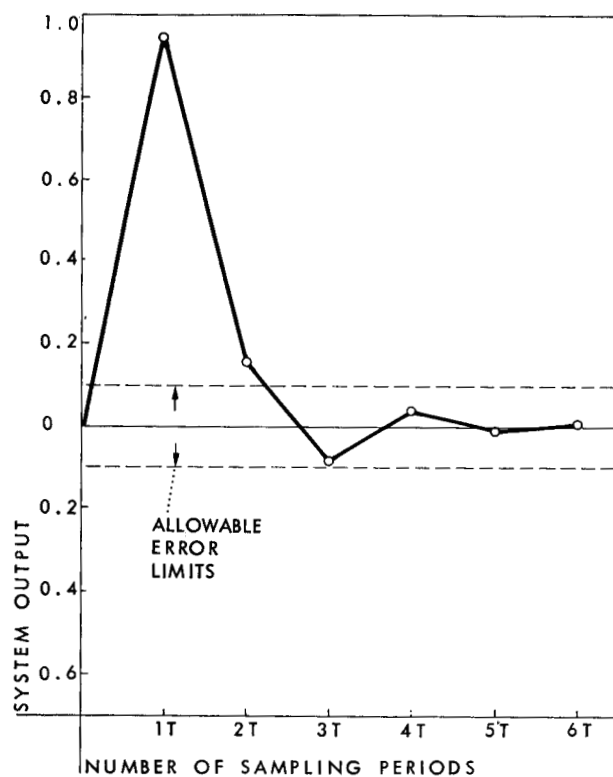
Figure 8 System response to step function input applied (at $t=0$).



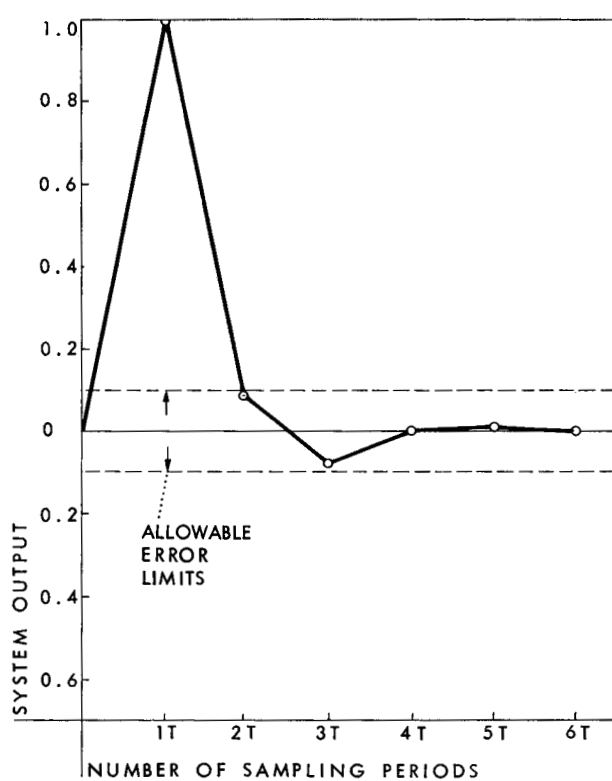
a) $T=600 \mu\text{sec}$, $T_a=540 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.



b) $T=5000 \mu\text{sec}$, $T_a=4000 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.



c) $T=600 \mu\text{sec}$, $T_a=180 \mu\text{sec}$, $K=1$, $a=R/L=5 \times 10^3$.



d) $T=600 \mu\text{sec}$, $T_a=540 \mu\text{sec}$, $K=1$, $a=R/L=40 \times 10^3$.

4. System operation

• Discussion of logical block diagram

The complete system implementation was considered from the standpoint of reliability, power, and space requirements. Hence, transistors were used exclusively as the active element. Wherever possible, moreover, silicon transistors were used to alleviate the problem of temperature dependency.

The logical block diagram in Fig. 11 shows the functional requirements of the system. Briefly, the system is intended to position sweep traces on a cathode-ray-tube presentation by comparing a voltage derived from the deflection-coil current to a reference voltage, sampling the resulting error before or during the sweep, and correcting the error before the next sweep is initiated. Required accuracy is one part in one thousand, and maximum correction per sweep is ten parts per thousand.

Operation can be described as beginning with the measurement of the current in each deflection coil as a voltage across the 10-ohm series resistor (1 unit error = 8 mv). The desired signal voltages across the resistors are identical in amplitude and waveform but 180° out of phase; the undesired signals due to power supply drift and ripple are also identical but in phase. Thus when the voltage across one resistor is inverted in the unity gain dc amplifier and summed with the voltage across the other at the input to the dc summing amplifier, the undesired signals cancel and the desired signals add, so that now unity error = 16 mv.

The deflection-coil signal to be inverted is selected so that there is 180° phase difference between the controlled output (deflection-coil current) and the summed deflection-coil signals. Because of this phase inversion, the input to the dc summing amplifier is the difference between any other dc amplifier inputs and the controlled output. These additional inputs are the controlling inputs, which in ppi mode provide a dc reference and which in sector-scan mode provide the sum of unresolved and resolved sweeps and the dc reference.

This difference between controlling input and controlled output is amplified 62.5 times in the dc amplifier and applied to the sampler, which at time T_s selects a one- μ sec sample. The voltage level at this point is one volt per unit error. The sampler has two parallel paths: one operates when the error is positive and the other when it is negative. If the error sample is positive, it is amplified by the emitter-follower-inverter combination and applied to the holding-circuit direction trigger to set the counter in a "count-down" condition. Originally, the amplified and inverted T_s pulse had set the trigger to the "count-up" condition, in which it would remain if the error sampled had been negative. In effect, then, the direction trigger preserves the polarity of the error.

The positive and negative error paths in the sampler are combined in a full-wave rectifier so that the sampler output is a positive pulse equal to the amplitude of the input error regardless of error polarity. This positive pulse is applied to an analog-to-digital converter consist-

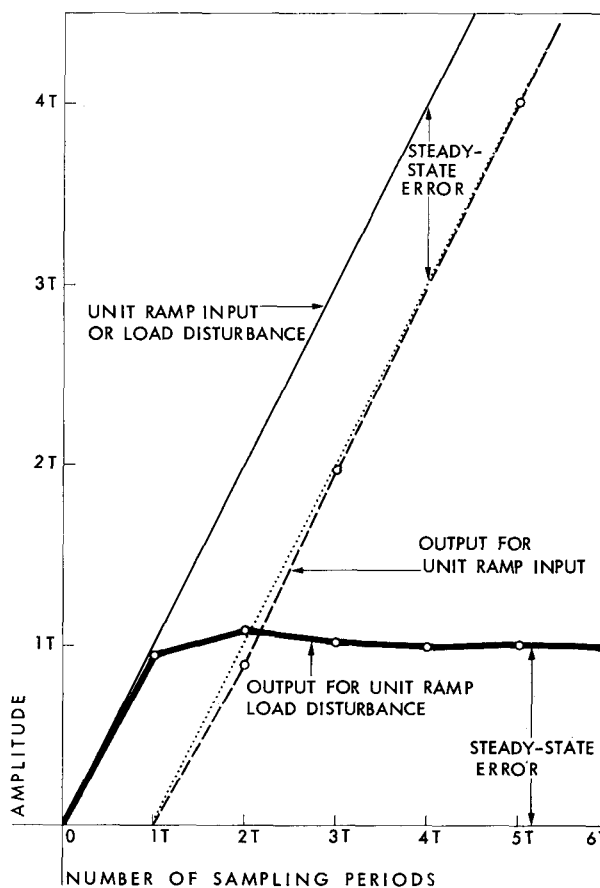


Figure 10 System response to unit ramp input and load disturbance. $\alpha=40,000$; $T=600$ μ sec; $T_d=540$ μ sec; $K=1$.

ing of the amplitude discriminator, write trigger and write gate. The amplitude discriminator converts the error sample into two negative pulses separated in time by the number of clock-pulse periods equal to the number of error units of the error sample. When the two negative pulses are applied to alternate sides of the write trigger, a 9-volt negative output equal in width to the separation of the triggering pulses results. The 9-volt trigger output is applied along with clock pulses to a negative AND circuit, which, together with an emitter-follower driver, make up the write gate. The number of clock pulses appearing at the write gate output represents the error sample amplitude.

From the write gate the clock pulses are applied through an OR circuit and emitter follower to the four-stage, binary-connected storage counter. The storage counter had previously been set to its count-up state by its direction trigger, which in turn was triggered by the T_s pulse. The storage counter thereby counts the clock pulses and stores them until desired. When the output of the write trigger returns to ground from -9 v, the positive shift flips the direction trigger, setting the storage counter to its count-down condition. Because the storage counter now has other than a zero count in it, the com-

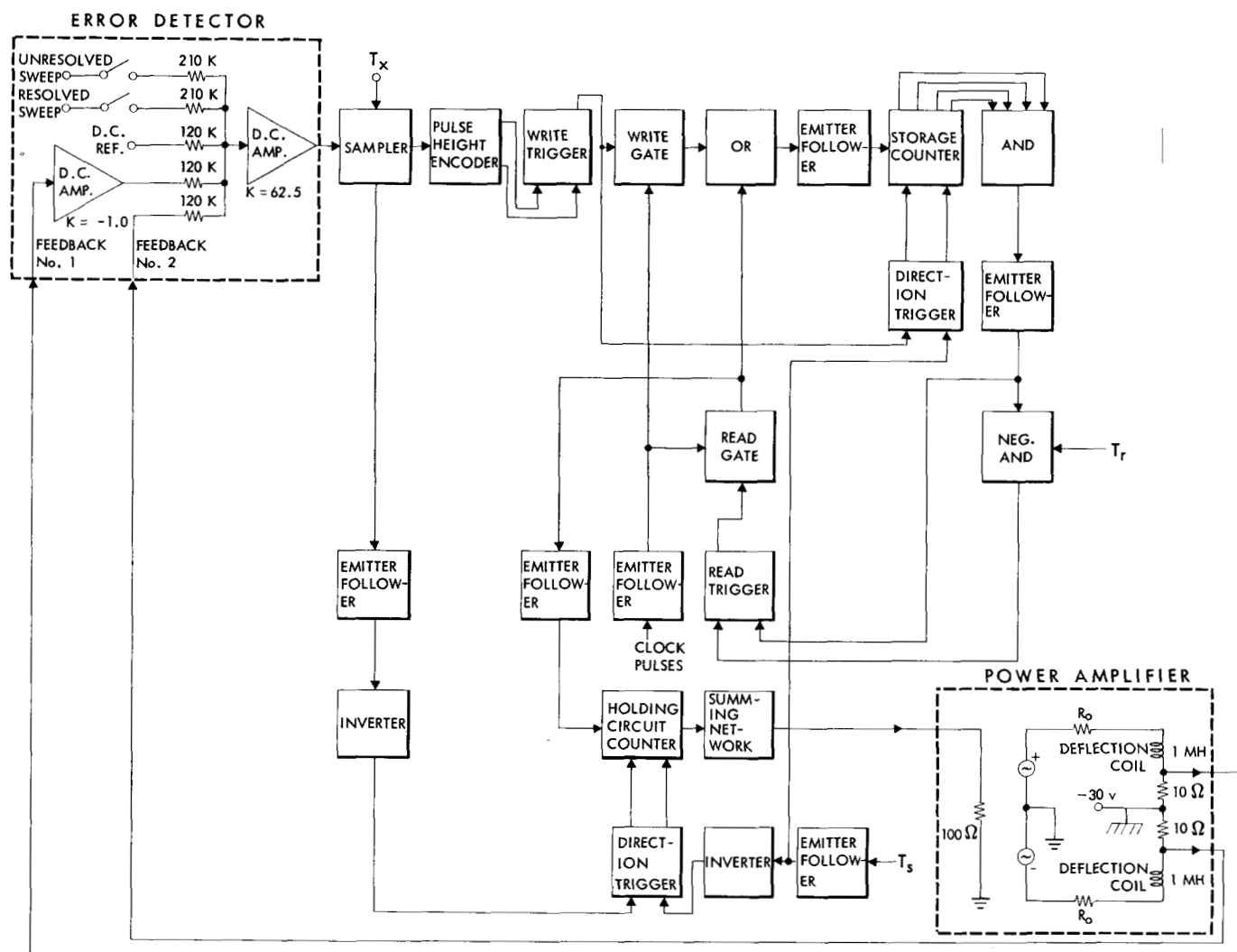
bination of AND circuit and emitter-follower connected to its four stages shifts in level from +9 v to ground. This shift, in turn, also shifts the level of the negative AND circuit from +9 v to ground.

At time T_r when correction is to be made for the error stored in the storage counter, the T_r pulse is applied to the other input of the negative AND circuit just mentioned. If there is an error signal, the count in the storage counter will have set both the AND and negative AND circuits to zero level. The positive T_r pulse can then pass through the negative AND circuit, flip the read trigger, and close the normally open read gate, allowing the signal to pass. If there is no error, the storage counter will remain with zero count in it and both the AND and negative AND circuits will be at +9 v. Since in this case, the T_r pulse cannot get through the negative AND to flip the read trigger, the read gate remains open for the cycle so that the signal cannot pass.

With the read gate closed, the clock pulses pass through

to the ten-stage binary-connected holding-circuit counter and cause it to alter its count up or down according to the condition of its direction trigger. Clock pulses are simultaneously applied through the OR circuit and emitter follower to the storage counter, which is now in its count-down condition. Each clock pulse that goes to the holding circuit also reduces the count in the storage counter until all storage counter triggers are in their zero state. At this time the AND circuit fed by the storage counter triggers shifts level from ground to +9 v and flips the read trigger off to open the read gate and stop the flow of clock pulses. In the process the holding circuit count is augmented or decreased (depending on error polarity) by the count stored in the storage counter, which represents the amplitude of the original error sample. The output current of the holding circuit summing network is changed proportionately. Deflection coil error is thereby corrected and the system is ready for another cycle.

Figure 11 System logical block diagram.



• Error detector and sampler

Figure 11 shows the summing network and amplifiers needed for error detection. The choice of two dc operational amplifiers over one difference amplifier was dictated by several factors. Inputs to the circuit vary from 8 mv to 6 v, and common mode rejection in a transistor difference amplifier is not characteristically good over such a wide dynamic range. In addition the summing network must work into a very low input impedance to achieve accuracy, and here again the input impedance to a difference amplifier is characteristically high. Finally, the problems of gain, bandwidth, and drift were easier to overcome in the operational amplifier design than in the difference amplifier design.

Operational amplifiers composed of three-stage feedback amplifiers that included drift-compensating circuits were selected. The amplifier used for inversion and level setting has unity gain. The summing amplifier has a gain of 62.5 and provides a one-volt output for a unit error input. Because the resolution of the system is ten unit errors per sweep, the active region of the summing amplifier output need be only ± 10 v for either positive or negative ten-unit errors.

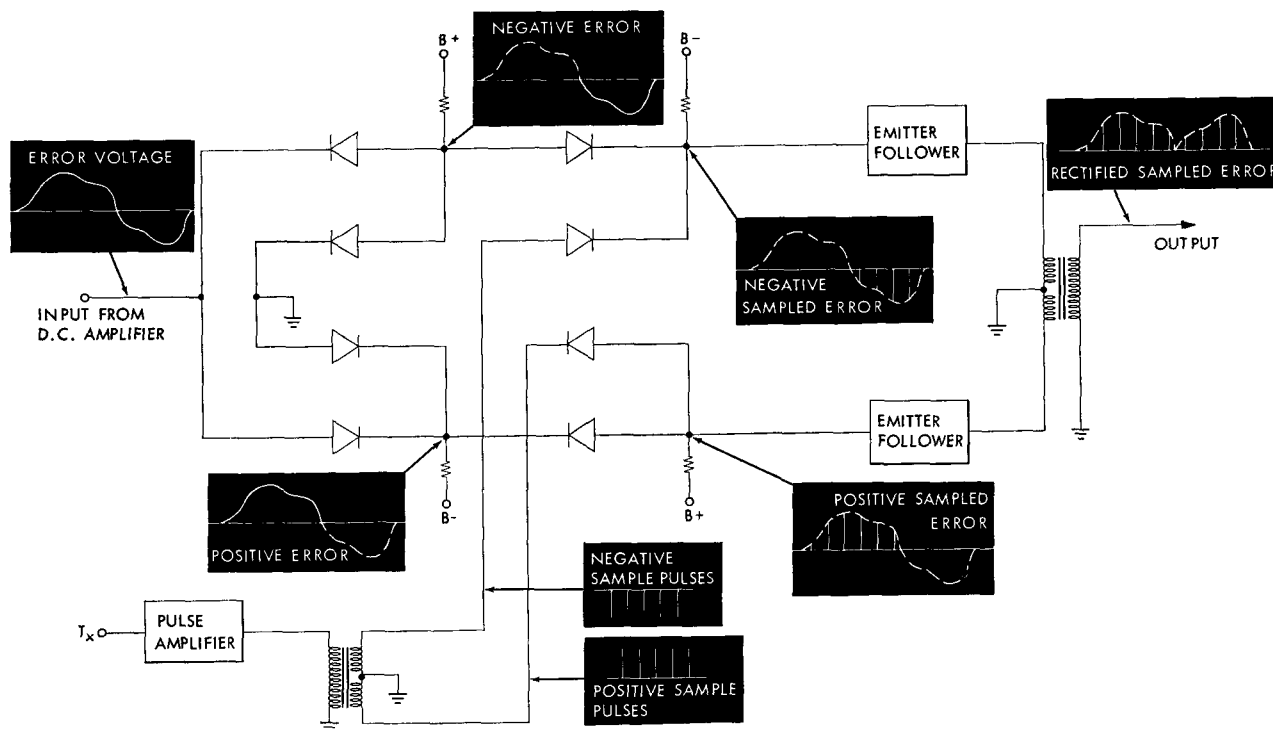
The continuous error signal from the dc amplifier is separated into positive and negative voltages by the first set of gates shown in Fig. 12, the sampler-circuit schematic. Each of these signals is applied to an input of a coincidence gate, the other input being a one- μ sec pulse which occurs at time T_x , and has a polarity corresponding to that of the error voltage. There is no output from the

gates unless both the error voltage and the sampling pulse are present simultaneously. When they are, the output from one gate or the other will be a one- μ sec pulse of amplitude and polarity corresponding to the error at T_x time. These pulses are then fed to emitter followers that drive a rectifying circuit, the output of which is a one- μ sec positive pulse of amplitude equal to the error regardless of input polarity. Because the sampler works at a relatively high level, the effects of level shift in the diode circuits are negligible. After this point the system is ac-coupled. The slight attenuation caused by the finite forward resistance of the diodes and less than unity gain in the emitter followers can be made up by adjusting the turns ratio of the pulse transformer at the output.

• Analog-to-digital converter

The analog-to-digital converter (see Fig. 13) consists of a rather novel combination of analog and digital circuits. Its purpose, which is to deliver a number of output digital clock pulses proportional to the amplitude of the input error pulse, is achieved as follows: Capacitor C is normally charged to a reference voltage much higher than the expected error voltage. When the error voltage is applied to the input, C charges to a voltage equal to the reference plus the error voltage and thereby reverses the bias on diodes D_1 and D_2 . The voltage on C then discharges toward ground through R until only the reference voltage is left, at which point the diodes once more begin conducting. Since the reference voltage is much

Figure 12 Error sampler.



larger than the maximum error voltage, only the linear portion of the exponential decay is used.

In opening and closing, diode D_2 alternately turns the regenerative amplifier on and off, causing an output pulse to appear at the transformer when C charges above reference voltage and again when it discharges to reference voltage. The interval Δt required to discharge C to the reference voltage is directly proportional to the amplitude of the error voltage. As a result the output pulses of the regenerative amplifier occur a time Δt apart. These output pulses, after being separated, are applied to a write trigger to produce a gating pulse of width Δt . Since the effect of this gating pulse at the negative AND circuit in the write gate is to allow clock pulses out for a time Δt , the direct relationship to error voltage is maintained. The circuit timing was designed so that one unit error (1 volt) corresponds to one clock pulse out.

• Storage and transfer circuits

The circuits diagrammed in Fig. 14 must receive the error input, store it for a specified time, and then transfer it to the holding-circuit counter. In addition, the storage counter must be reset to zero count and ready for the next cycle when the transfer is accomplished. A four-stage, bidirectional, binary-connected counter is used.

Because the maximum error is ten units (or ten clock pulses), four stages of triggers can provide the necessary storage capacity. The direction trigger controls the gates which allow either the right sides or the left sides of the triggers to be connected in binary. Thus, the counter will

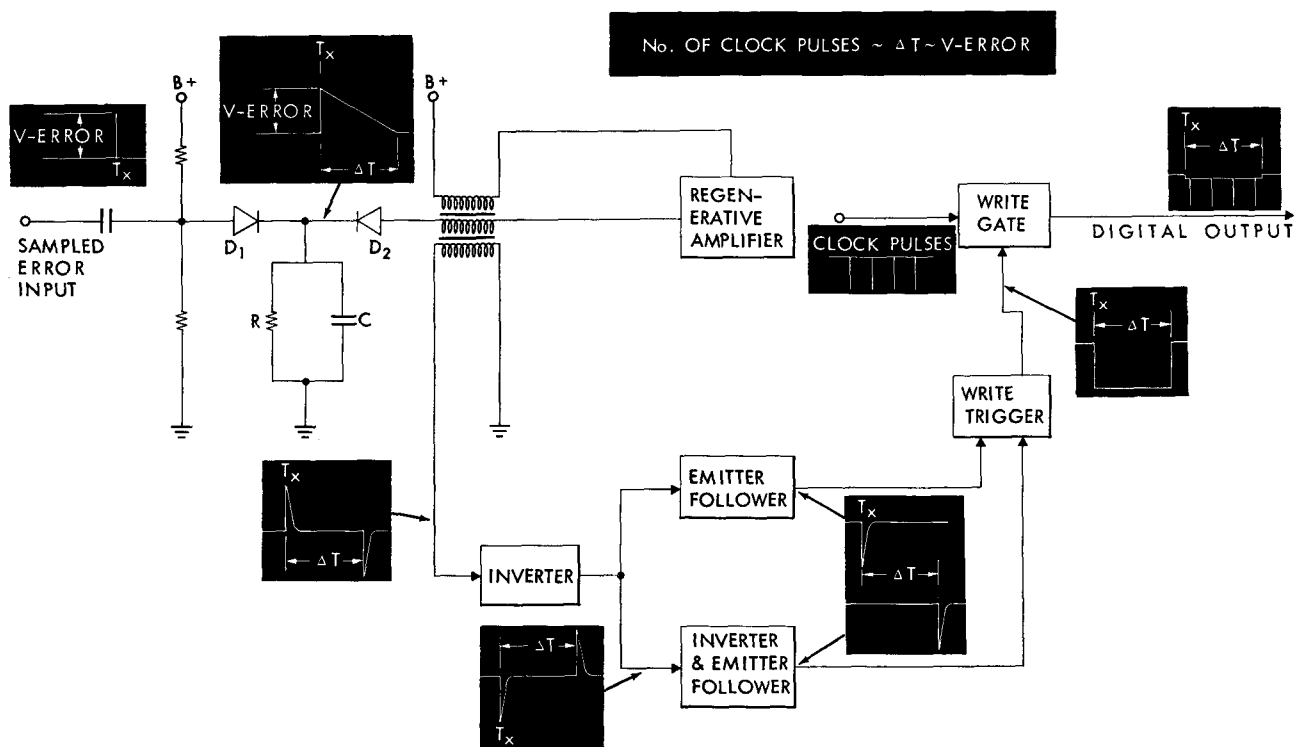
count up when the right sides are connected and will count down when the left sides are connected. The gates are of the diode-capacitor type, since ac coupling is used throughout the digital portion of the system. The triggers are a binary-connected standard saturating circuit using Texas Instrument 904A *npn* transistors. The repetition rate (clock frequency) used in the system is approximately 160 kc, although the counter itself will operate up to about 500 kc.

• Digital-to-analog converter and compensator

The digital-to-analog converter comprises the last group of circuits and includes the holding circuit counter and summing network (ladder network decoder). It must convert the digital error transferred from the storage counter to an analog pulse height, integrate this pulse height arithmetically and hold or clamp the output between sampling instants. The circuit is shown in detail in Fig. 15. Because there are ten stages, the summing network output may represent 2^{10} , or 1024, discrete current levels, which is more than sufficient to guarantee the 0.1% precision requirement. The holding circuit counter itself is a replica of the storage counter but has six additional stages. It operates in exactly the same manner and at the same repetition rate.

Overall converter action is indicated by the waveforms in Fig. 16. The two input cycles represent two successive negative errors of four and six units in digital form. The output waveform demonstrates the conversion of these errors to analog pulses. These are represented by the

Figure 13 Analog-to-digital converter.



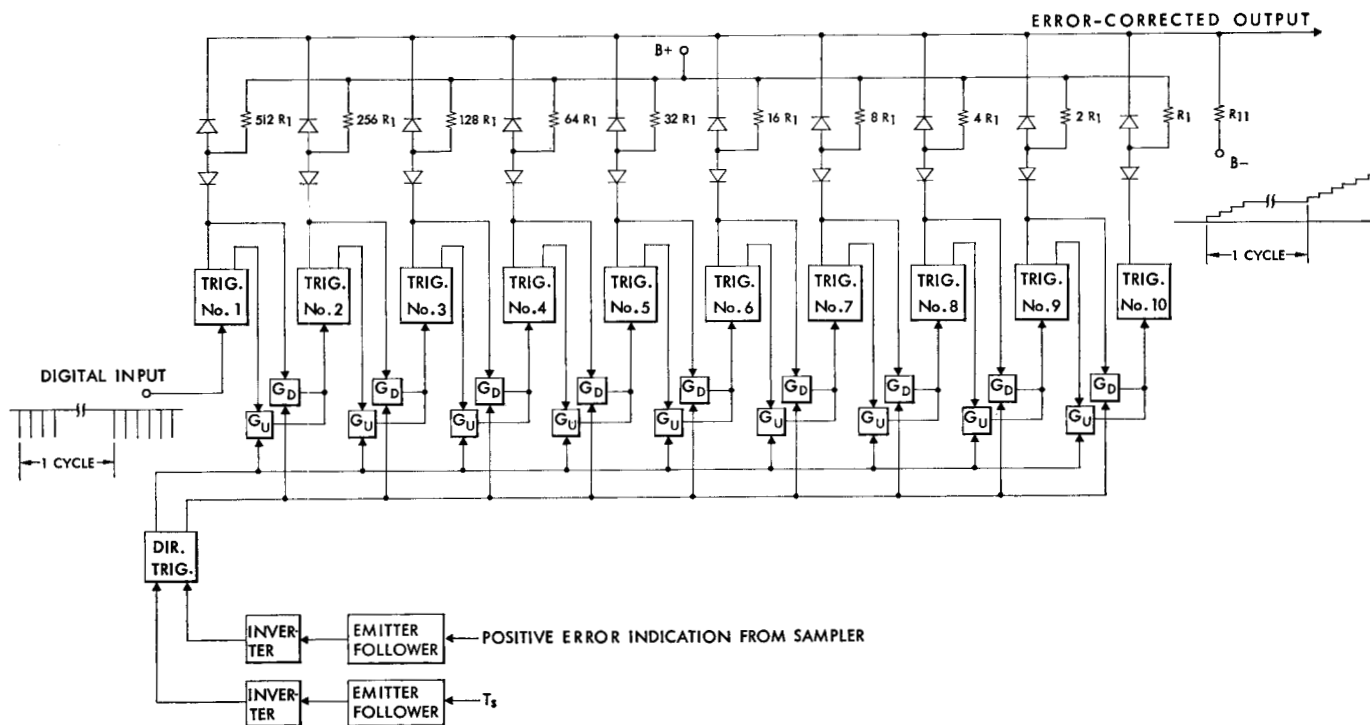
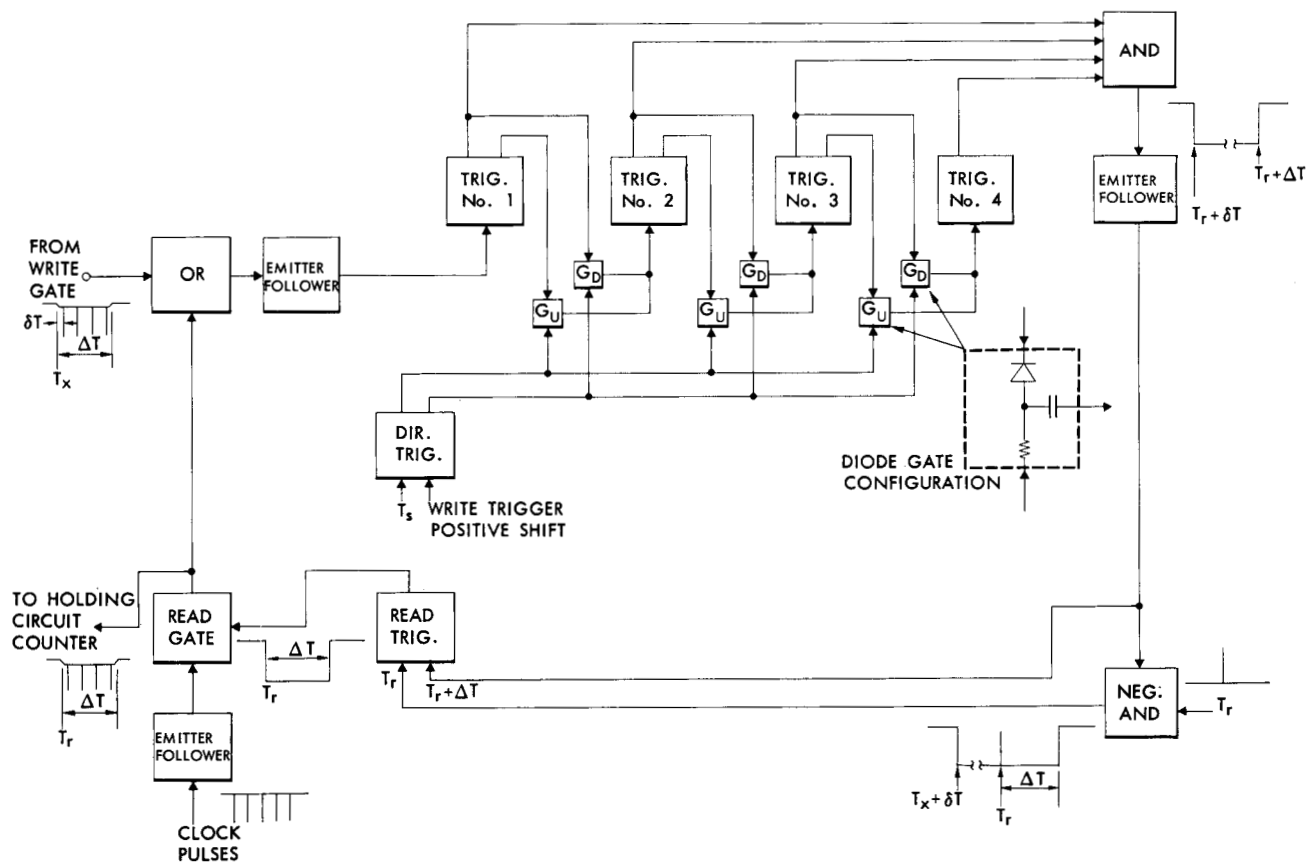


Figure 14 Storage counter and transfer circuits.

Figure 15 Digital-to-analog converter and compensator.



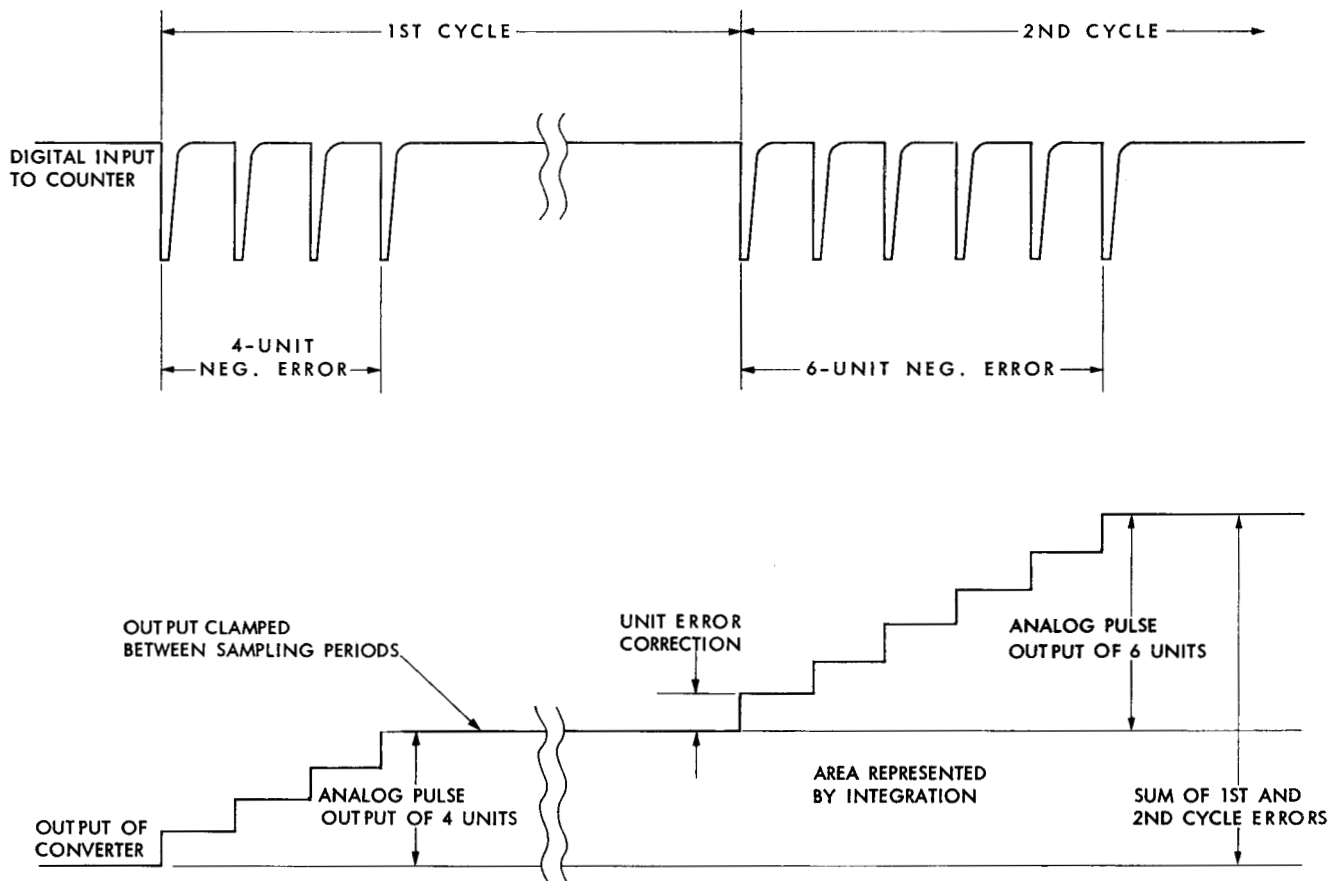


Figure 16 Digital-to-analog converter waveforms.

Figure 17a Top: Input sawtooth, duration 600 μ sec, maximum voltage 10v.
Bottom: Closed loop output.

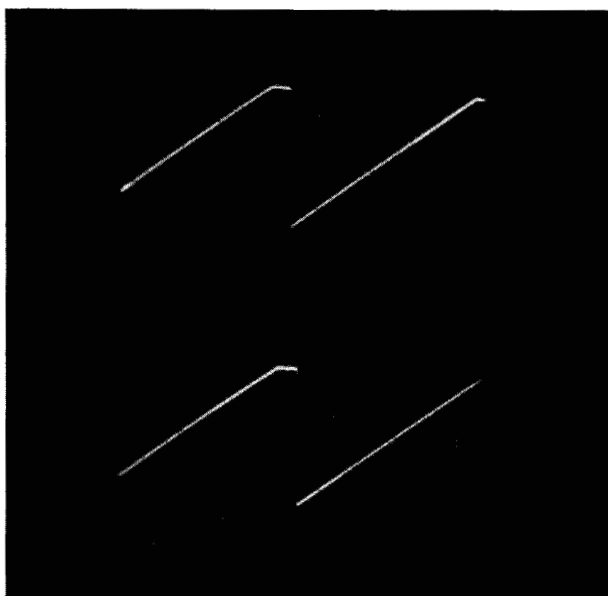
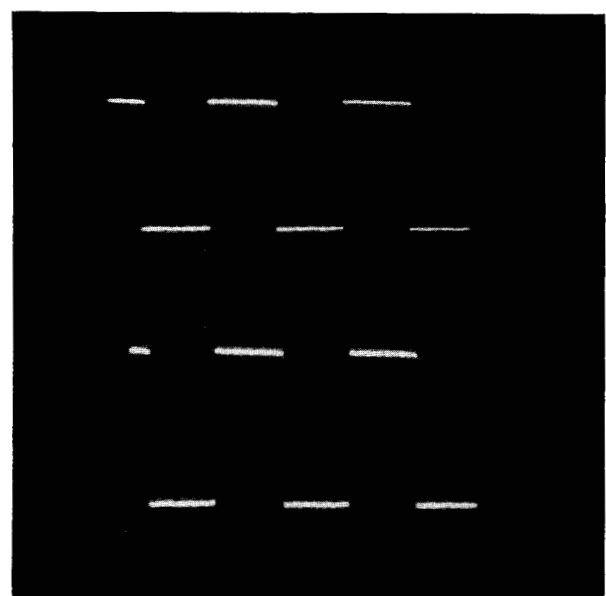


Figure 17b Top: Square wave input at 1 cps.
Bottom: Closed loop output.



large steps. Because the output is held, or clamped, to a particular level between sampling instants, the reconstructed analog pulse, per se, does not appear. The arithmetic integration is obvious in that each analog error is being added sequentially to the sum of all previous errors. The effect of the circuit, then, is to convert the digital error to a change in dc level in the load equal and opposite to the input error of the system.

● System performance

Open and closed loop tests on the dc restoration system demonstrate that performance is as specified. In essence, the specification requires that the system correct for a step-function disturbance of less than 10 units within one sampling period and to an accuracy of plus or minus one unit error. The system meets these requirements quite adequately, as predicted. There is ripple in the load contributed by the deflection amplifier which the system cannot correct. However, this ripple is less than one unit error in amplitude.

Oscillographs illustrating system performance are shown in Figs. 17 to 19. The fastest sampling rate of 1600 cps ($600 \mu\text{sec}$) was used for the tests in order to prove performance under the most stringent conditions.

Figure 17a shows the system following a slow sawtooth waveform. This is indicative of its response to a ramp or constant-velocity type input. As predicted, the system follows with no difficulty. Figure 17b illustrates system response to a low-frequency square wave which can be considered as a step-function input. The output will fol-

low the input in one sampling period if the input amplitude is limited to 10 unit errors (above which the system saturates). The oscillograph actually represents an input amplitude of approximately 120 unit errors. For this condition the system must take $120/10$, or 12, sampling periods to rise to the input amplitude. This can be seen graphically in Fig. 17c, where response for a 20-cps square wave input with a 120-unit error amplitude is shown. Twelve steps representing the twelve sampling periods can actually be counted. The rise time was almost exactly 12 times the $600\text{-}\mu\text{sec}$ sampling interval, or 7.2 msec.

Figure 17d illustrates the system response to a similar input at a 20-cps sine wave. Note that, as with the square wave, the steps become more apparent with increasing frequency because of the decreasing number of samples taken per cycle. Fig. 18 provides an interesting demonstration of how the system corrects for load disturbances. With the loop open, a 5-cps sine wave applied to the deflection amplifier results in an output across the sampling resistors, as can be seen at the top of the oscillograph. Closing the loop with the same signal applied reduces the disturbance to that shown in the bottom of the oscillograph. The 5-cps disturbance is barely apparent here, having been reduced by a factor of about 25 to less than one unit error. This effect is even more pronounced at lower frequencies.

Figure 17c **Top: Square wave input at 20 cps.**
Bottom: Closed loop output.

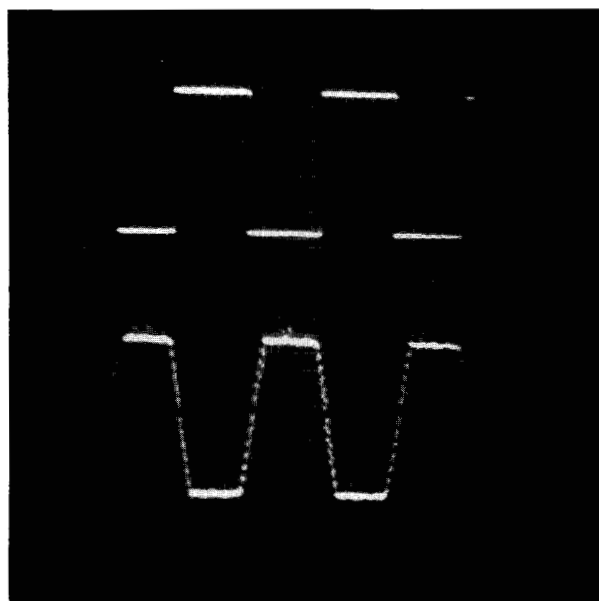
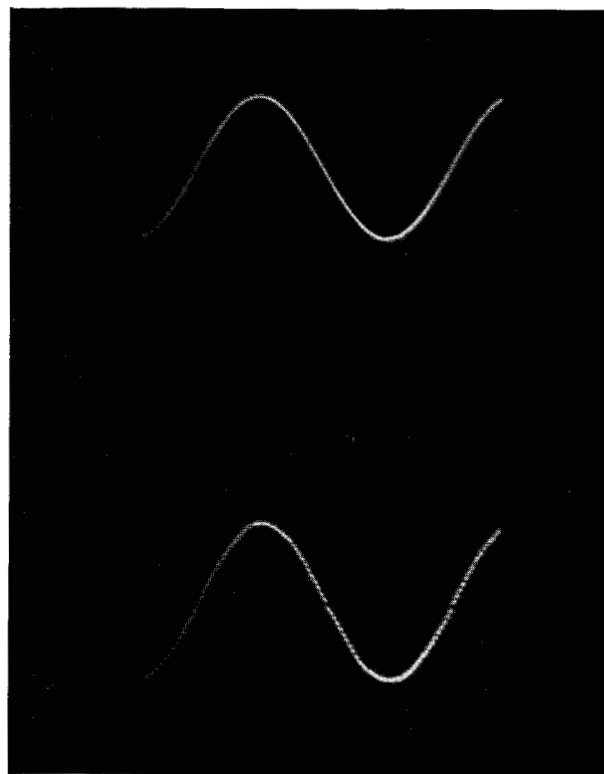


Figure 17d **Top: 20 cps sine wave input.**
Bottom: Closed loop output.



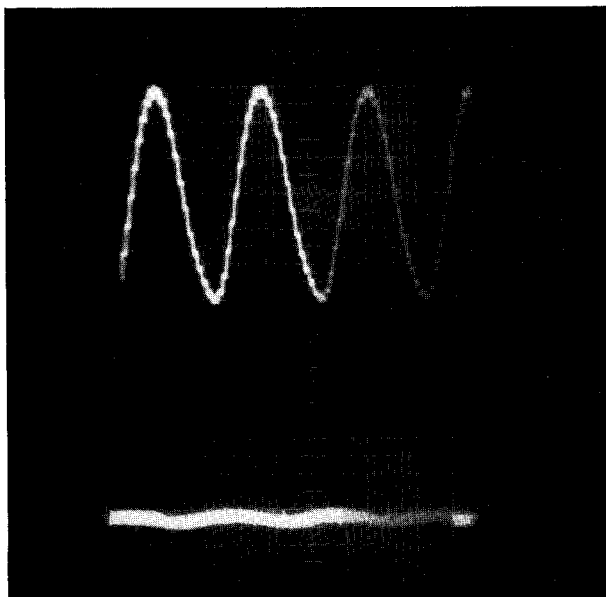


Figure 18 Top: Open loop output with 5 cps sine wave disturbance applied to power amplifier.
Bottom: Closed loop output under same condition.

Fig. 19 illustrates the performance of the system under conditions similar to those for sector-scan mode. A 500- μ sec sweep waveform (a) is applied to the unresolved sweep input of the dc summing amplifier. This sweep is synchronized with system triggers so that error is sampled each time at a fixed instant on the sweep. The voltage level of the amplified sweep waveform (b) is sampled by application of the sampling pulse (c), and the system output is then a dc level (d) corresponding to the voltage of the sweep at sampling time (c).

The restoration system operated open-loop acts as an almost perfect integrator, a characteristic which, as has been explained, is due to the action of the digital-to-analog converter. A sine wave introduced at the input produces a sine wave at the output phase-shifted 90° . A square-wave input produces a triangular-wave output, as would be expected.

5. Conclusions

The application of sampled-data theory to the synthesis of a feedback control system and its implementation have both been described in this paper. Detailed tests have indicated that the system corrects for a step-function disturbance of less than 10 unit errors within one sampling period and to an accuracy of plus or minus one unit error, which adequately fulfills the specifications. The advantages of digital compensation must, of course, be weighed against the increase in equipment and cost over standard methods of compensation. However, the high accuracy and flexibility which result from digitally-compensated control makes it desirable for many applications.

The z-transform methods used in the design of this system are applicable to an entire class of control system problems where it is necessary or desirable to utilize digital-computer techniques for system compensation. It is also expected that in view of the increasing interest in utilizing digital computers in industrial process control systems, the z-transform methods will become a powerful operational tool for the control engineer.

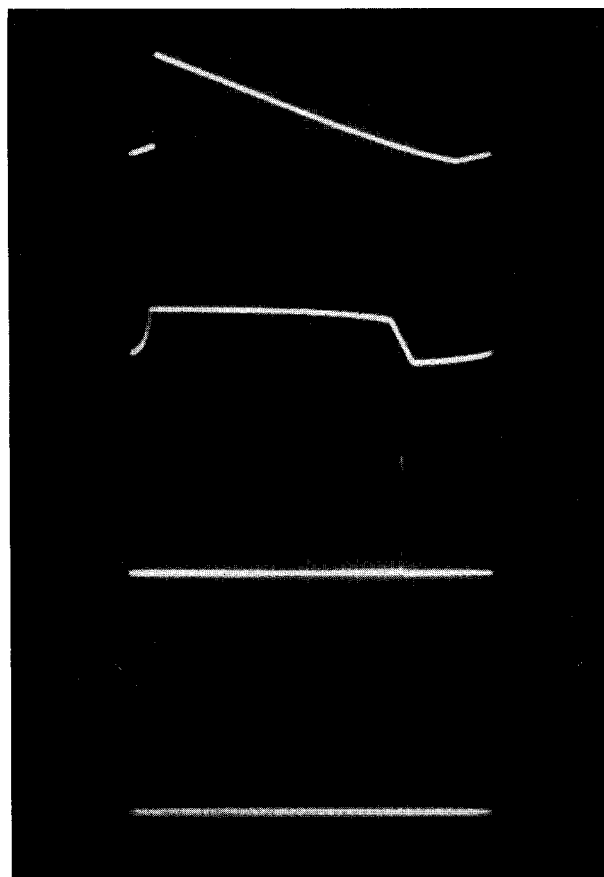
Acknowledgments

The work reported on in this paper was supported by the U. S. Air Force under Contract Number AF 33(038) 20147.

The authors wish to acknowledge the encouragement and assistance given them by D. R. Baldauf, J. L. Craft, R. A. Henle, J. C. Logue and J. L. Sweeney. Thanks are also due A. C. Farrell and R. G. Simmons for their assistance in constructing the breadboard model and to Miss K. M. Clark for her cooperation in preparing the manuscript.

Figure 19 From top to bottom:

- 500 μ sec sawtooth applied to unresolved sweep input of system.
- Output of dc summing amplifier showing it coming out of saturation at time T_x .
- T_x sampling pulse.
- Dc level of system output shown over several sweep periods.



List of symbols

t = Time variable	$G(z)$ = System transfer function as function of z
$s \triangleq$ Complex variable $\sigma + j\omega$	$E(z) = R(z) - C(z)$ = System error as function of z
$z \triangleq$ Transform variable e^{sT}	$H(z)$ = Feedback transfer function as function of z
T_s = System trigger	$F(z)$ = General function of z
T_r = Sweep recovery time	$\mathcal{L}$ = Laplace transform operator
$T = T_x$ = Sampling time	$\mathcal{Z}$ = z -transform operator
$T_d = T_r - T_x$ = System transportation lag (delay time)	e_{ss} = Steady-state error
$\Delta T = T - T_d$	$K \triangleq$ Effective loop voltage gain
a = Inverse of load time constant	K_c = Continuous system loop gain
$r(t)$ = Input as function of time	K_s = Sampled system loop gain
$c(t)$ = Output or response as function of time	K_v = Velocity error constant
$g(t)$ = System transfer function as function of time	$C(nT)$ = Output or response at sampling instants ($n = 0, 1, 2, 3 \dots$)
$e(t) = r(t) - c(t)$ = System error as function of time	ω_0 = Undamped natural frequency of the system in radians
$h(t)$ = Feedback transfer function as function of time	ζ = Damping ratio
$R(z)$ = Input signal as function of z	σ = Damping factor
$C(z)$ = Output signal or response as function of z	

References

1. Barker, R. H., "The Pulse Transfer Function and Its Application to Sampling Servo Systems," *Proceedings IEE*, London, England, **99**, Pt. IV, 302-17 (December, 1952).
2. Linvill, W. K., "Sampled-Data Control Systems Studied Through Comparison of Sampling with Amplitude Modulation," *Transactions AIEE*, **70**, Pt. II, 1779-99 (1951).
3. Ragazzini, J. R. and Zadeh, L. A., "Analysis of Sampled-Data Systems," *Transactions AIEE*, **71**, Pt. II, 225-34 (1952).
4. Truxal, J. G., *Control System Synthesis*, McGraw-Hill, New York, 1955; Chapter 9.
5. Johnson, G. W., Lindorf, D. P. and Nordling, C. G., "Extension of Continuous-Data System Design Techniques to Sampled-Data Control Systems," *Transactions AIEE*, **174**, Pt. II, 252-63 (September, 1955).
6. James, H. M., Nichols, N. B. and Phillips, R. S.; *Theory of Servomechanisms*, McGraw-Hill, New York, 1947.
7. Sklansky, J., "Network Compensation of Error-Sampled Control Systems," Technical Report T-7/B, Electronic Research Laboratories, Columbia University, New York, April 1955.
8. Brown, G. S. and Campbell, D. P.; *Principles of Servomechanisms*, John Wiley and Sons, New York, 1948.
9. Knapp, C. H., Shapiro, E. and Thorpe, R. A.; "The z -Transform: A Method for Analysis of Sampled-Data Systems," Research Report RC-17, IBM Research Center, Poughkeepsie, New York, December 1957. (Copies are available on request to the Editor.)
10. Chu, Yaohan, "Feedback Control Systems with Dead-Time Lag or Distributed Lag by Root-Locus Method," *Transactions AIEE*, **71**, Pt. II, 291-96 (November 1952).

Revised manuscript received September 12, 1957